

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: March 2, 2006, 06:32:30 ; Search time 1152 Seconds

(without alignments)
799.304 Million cell updates/sec

Title: US-10-041-030-4

Perfect score: 2290

Sequence: 1 MRSPPGEHCAPKPEPVXYG.....ATQLVGEQNCIKLIPGSPID 420

Scoring table:
BLOSUM62
Xgapop 10.0 , Xgapext 0.5
Ygapop 10.0 , Ygapext 0.5
Fgapop 6.0 , Fgapext 7.0
Delop 6.0 , Delext 7.0

Searched: 7218512 segs, 109618870 residues

Total number of hits satisfying chosen parameters: 14437024

Minimum DB seg length: 0
Maximum DB seg length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame_p2n.model -DEV=xlh
-Q=/abs/ABSSMB.spool/US10041030/runat_01032006_134413_22030/app_query.fasta.1
-DB=Published Applications NA New -QMT=fastcap -SURF=p2n.rnpbn -MINMATCH=0.1
-LOOPEX=0 -LOOPEXT=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62
-TRANS=humand.cdi -LIST=45 -POCALIGN=200 -THR SCORE=pct -THR MAX=100
-THR MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext -HEADSIZE=500 -MINLEN=0
-MAXLEN=200000000 -HOST=abs05h
-USER=US10041030 @CEN 1.1 220 @runat 01032006 134413 22030 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSPELOCK=100 -LONGLOC -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THRBADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA New:

1: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
2: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US11_NEW_PUB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|--------|--------------------|-------|--|
| 1 | 1482.5 | 64.7 | 2508 | 9 US-11-072-512-1958 Sequence 1958, Ap |
| 2 | 711.5 | 31.1 | 1105 | 8 US-10-750-185-40960 Sequence 40960, A |
| 3 | 711.5 | 31.1 | 1105 | 8 US-10-750-623-40960 Sequence 40960, A |
| 4 | 293 | 12.8 | 1345 | 6 US-09-925-065A-679103 Sequence 679103, |

| 5 | 102 | 4.5 | 3078 | 8 | US-10-750-185-38950 | Sequence 38950, A |
|----|-------|-----|--------|--|-----------------------|-------------------|
| 6 | 102 | 4.5 | 3078 | 8 <td>US-10-750-623-38950</td> <td>Sequence 38950, A</td> | US-10-750-623-38950 | Sequence 38950, A |
| 7 | 101.5 | 4.4 | 11916 | 12 <td>US-11-121-086-12</td> <td>Sequence 12, Appl</td> | US-11-121-086-12 | Sequence 12, Appl |
| 8 | 101.5 | 4.4 | 16781 | 12 <td>US-11-121-086-14</td> <td>Sequence 14, Appl</td> | US-11-121-086-14 | Sequence 14, Appl |
| 9 | 100.5 | 4.4 | 1664 | 12 <td>US-11-013-247A-16</td> <td>Sequence 16, Appl</td> | US-11-013-247A-16 | Sequence 16, Appl |
| 10 | 97.5 | 4.3 | 164810 | 12 <td>US-11-121-086-4</td> <td>Sequence 4, Appl</td> | US-11-121-086-4 | Sequence 4, Appl |
| 11 | 97 | 4.2 | 21728 | 7 <td>US-10-330-773-362</td> <td>Sequence 362, App</td> | US-10-330-773-362 | Sequence 362, App |
| 12 | 97 | 4.2 | 150468 | 12 <td>US-11-112-908-56</td> <td>Sequence 56, Appl</td> | US-11-112-908-56 | Sequence 56, Appl |
| 13 | 97 | 4.2 | 193789 | 12 <td>US-11-112-908-55</td> <td>Sequence 55, Appl</td> | US-11-112-908-55 | Sequence 55, Appl |
| 14 | 96.5 | 4.2 | 1770 | 12 <td>US-11-183-131-41</td> <td>Sequence 41, Appl</td> | US-11-183-131-41 | Sequence 41, Appl |
| 15 | 94.5 | 4.1 | 1363 | 8 <td>US-10-750-185-32846</td> <td>Sequence 32846, A</td> | US-10-750-185-32846 | Sequence 32846, A |
| 16 | 94.5 | 4.1 | 1363 | 8 <td>US-10-750-623-32846</td> <td>Sequence 32846, A</td> | US-10-750-623-32846 | Sequence 32846, A |
| 17 | 94.5 | 4.1 | 3510 | 9 <td>US-11-114-962-9</td> <td>Sequence 9, Appl</td> | US-11-114-962-9 | Sequence 9, Appl |
| 18 | 94 | 4.1 | 3430 | 8 <td>US-10-750-185-62790</td> <td>Sequence 62790, A</td> | US-10-750-185-62790 | Sequence 62790, A |
| 19 | 94 | 4.1 | 3430 | 8 <td>US-10-750-623-62790</td> <td>Sequence 62790, A</td> | US-10-750-623-62790 | Sequence 62790, A |
| 20 | 94 | 4.1 | 53916 | 7 <td>US-10-330-773-65</td> <td>Sequence 65, Appl</td> | US-10-330-773-65 | Sequence 65, Appl |
| 21 | 93.5 | 4.1 | 1400 | 12 <td>US-11-136-527-5713</td> <td>Sequence 5713, Ap</td> | US-11-136-527-5713 | Sequence 5713, Ap |
| 22 | 93.5 | 4.1 | 1650 | 12 <td>US-11-013-247A-1</td> <td>Sequence 1, Appl</td> | US-11-013-247A-1 | Sequence 1, Appl |
| 23 | 93.5 | 4.1 | 2475 | 12 <td>US-11-136-527-1617</td> <td>Sequence 1617, Ap</td> | US-11-136-527-1617 | Sequence 1617, Ap |
| 24 | 93.5 | 4.1 | 6510 | 8 <td>US-10-995-561-112</td> <td>Sequence 112, App</td> | US-10-995-561-112 | Sequence 112, App |
| 25 | 93.5 | 4.1 | 7823 | 8 <td>US-10-995-561-117</td> <td>Sequence 117, App</td> | US-10-995-561-117 | Sequence 117, App |
| 26 | 93 | 4.1 | 1224 | 8 <td>US-10-750-185-45251</td> <td>Sequence 45251, A</td> | US-10-750-185-45251 | Sequence 45251, A |
| 27 | 93 | 4.1 | 1224 | 8 <td>US-10-750-623-45251</td> <td>Sequence 45251, A</td> | US-10-750-623-45251 | Sequence 45251, A |
| 28 | 92.5 | 4.0 | 590 | 6 <td>US-09-925-065A-458917</td> <td>Sequence 458917,</td> | US-09-925-065A-458917 | Sequence 458917, |
| 29 | 92 | 4.0 | 1560 | 7 <td>US-10-787-906-12</td> <td>Sequence 12, Appl</td> | US-10-787-906-12 | Sequence 12, Appl |
| 30 | 92 | 4.0 | 2483 | 9 <td>US-11-072-512-1600</td> <td>Sequence 1600, Ap</td> | US-11-072-512-1600 | Sequence 1600, Ap |
| 31 | 92 | 4.0 | 2767 | 8 <td>US-10-821-234-43</td> <td>Sequence 43, Appl</td> | US-10-821-234-43 | Sequence 43, Appl |
| 32 | 92 | 4.0 | 2936 | 12 <td>US-11-136-527-2243</td> <td>Sequence 2243, Ap</td> | US-11-136-527-2243 | Sequence 2243, Ap |
| 33 | 92 | 4.0 | 5928 | 8 <td>US-10-863-093-25</td> <td>Sequence 25, Appl</td> | US-10-863-093-25 | Sequence 25, Appl |
| 34 | 91.5 | 4.0 | 1569 | 12 <td>US-11-076-427A-29</td> <td>Sequence 29, Appl</td> | US-11-076-427A-29 | Sequence 29, Appl |
| 35 | 91.5 | 4.0 | 1571 | 12 <td>US-11-076-427A-27</td> <td>Sequence 27, Appl</td> | US-11-076-427A-27 | Sequence 27, Appl |
| 36 | 91.5 | 4.0 | 7893 | 12 <td>US-11-186-731-3</td> <td>Sequence 3, Appl</td> | US-11-186-731-3 | Sequence 3, Appl |
| 37 | 91.5 | 4.0 | 8106 | 12 <td>US-11-186-731-1</td> <td>Sequence 1, Appl</td> | US-11-186-731-1 | Sequence 1, Appl |
| 38 | 91.5 | 4.0 | 8250 | 8 <td>US-10-453-372-769</td> <td>Sequence 769, App</td> | US-10-453-372-769 | Sequence 769, App |
| 39 | 91.5 | 4.0 | 23907 | 12 <td>US-11-186-731-6</td> <td>Sequence 6, Appl</td> | US-11-186-731-6 | Sequence 6, Appl |
| 40 | 91.5 | 4.0 | 24120 | 12 <td>US-11-186-731-4</td> <td>Sequence 4, Appl</td> | US-11-186-731-4 | Sequence 4, Appl |
| 41 | 91.5 | 4.0 | 162289 | 12 <td>US-11-121-086-20</td> <td>Sequence 20, Appl</td> | US-11-121-086-20 | Sequence 20, Appl |
| 42 | 91 | 4.0 | 2156 | 12 <td>US-11-136-527-2936</td> <td>Sequence 2936, Ap</td> | US-11-136-527-2936 | Sequence 2936, Ap |
| 43 | 90.5 | 4.0 | 2736 | 9 <td>US-11-072-512-258</td> <td>Sequence 258, App</td> | US-11-072-512-258 | Sequence 258, App |
| 44 | 90 | 3.9 | 1970 | 6 <td>US-09-925-065A-551509</td> <td>Sequence 551509,</td> | US-09-925-065A-551509 | Sequence 551509, |
| 45 | 90 | 3.9 | 1970 | 6 <td>US-09-925-065A-551510</td> <td>Sequence 551510,</td> | US-09-925-065A-551510 | Sequence 551510, |

ALIGNMENTS

RESULT 1
US-11-072-512-1958
Sequence 1958, Application US/11072512
Publication No. US20060029945A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHITO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOMYUKI
APPLICANT: NAGAHARA, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: Novel full length cDNA
FILE REFERENCE: 084335-0191
CURRENT APPLICATION NUMBER: US/11/072, 512
CURRENT FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: US 60/350, 978

```

; PRIOR FILING DATE: 2002-01-25
; PRIOR APPLICATION NUMBER: JP 2001-379298
; PRIOR FILING DATE: 2001-11-05
; NUMBER OF SEQ ID NOS: 4096
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1958
; LENGTH: 2508
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-072-512-1958

Alignment Scores:
Pred. No.: 5,33e-127 Length: 2508
Score: 1482.50 Matches: 275
Percent Similarity: 74.2% Conservative: 36
Best Local Similarity: 65.6% Mismatches: 65
Query Match: 64.7% Indels: 43
DB: Gaps: 4

US-10-041-030-4 (1-420) x US-11-072-512-1958 (1-2508)

QY 3 SerProGlyGlnGluGlnHisCysAlaProAsnLysGluProVallyrGlyGluLeu 22
DB 208 TCTCCCGCT-----GAAGATGCGCAGCCAGCGAGAGCCCATCAAGTATGTGAATC 261
QY 23 ValValLeuGlyrYrAsnGlyAlaLeuProAsnGlyAspArgLysrArg 42
DB 261 ----- 261
QY 43 PheAlaLeuTyrylAspArgProLysAlaAsnGlyVallyrProSerThrValHisValIle 62
DB 262 -----ATC 264
QY 63 SerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSerTyThr 82
DB 265 TCACAGCGCGCTCTCTCCAGAGCAGCTAGTAACGGTGTGACAGCAGCATCTCGTATACA 324
QY 83 LeuSerArgAsnGlnThrValValValGlyrYrThrHisAspLysAspThrAspMetPhe 102
DB 325 CTGTCCCGGAGCCACTCGCTCATAGTGAATATACATGATAGCCGACACAGACATTTTC 384
QY 103 GlnValGlyrArgSerThrGlnSerProLysAspPheValValThrAspThrIleSerGly 122
DB 385 CAGATTGCGCGCTCCACAGAGACATGATGATCTGTGGTAACAGACAGTCCCGCGGA 444
QY 123 SerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArg 142
DB 445 GGA---GGGCGTGGCCGAGCGCCCTTCTGCCAGAGCAACAATCTCCCTATGCTCGCGC 501
QY 143 IleValCysAspArgAsnGluProTyThrAlaArgLysPheAlaIleAspSer 162
DB 502 ATCTCTGTATACCGCGCCGACCTATATAGCCCGCATCTATGCGGTGCTTGATGTC 561
QY 163 SerLysAsnLysPheLeuGlyGluValAlaAlaLysTrpLysAsnProAspArgLysIleMet 182
DB 562 TCTAGCAACATCTTCTTGTGAGAGCGAGCGCCAAATGCGGAGCCCAAGTGGCTATG 621
QY 183 AspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThrGlnGlu 202
DB 622 GATGAGATGACCAACAATGAGTCTGTGATGACCCGCGAGCGGCTTCTCCGAGAGC 681
QY 203 SerGlnProGlyValLTPArgGluLysSerValCysGlyLysAspValTyThrIleAsnGlu 222
DB 682 TCAAGCCCGGAGTGTCTGGCGGAGATCTCGTCTGTGGAAATGTGTACACATTCGGGAG 741
QY 223 ThrArgSerAlaGlnGlnArgGlyLysLeuValGlnSerGluThrAsnValLeuGlnAsp 242
DB 742 AGCCGCTCAACCCAGCGAGCGGCGCAAGCTGTGTGAAAACAGTCAACGTCCTCAAGAC 801
QY 243 GlySerLeuLysAspLeuCysGlyAlaThrLeuLeuTrpArgThrAlaAspGlyLeuPhe 262
DB 802 GGCTCTCTATCACTGATGTGTGGGCGACACTGCTGTGGCGCACACCGCGGGGCTGCTG 861

```

```

QY 263 HisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAlaIleArgPro 282
DB 862 CGGGCTCCACACTGAAGCAACTGAGAGCGCCAGCGGACAGGAAATGACAGCGCGGCC 921
QY 283 GlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGluVal 302
DB 922 CAGTGGCCCGTGGGCTTGGCACTTGGCTTGGCCAGCCAGCCCGGTGGCGCACAGCG 961
QY 303 GlnGluLys---GlnProTrpAlaTyLeuSerCysGlyYrHisValHisGlyYrHisAsn 321
DB 982 CCGGACAAACAGAGCCCTTGGCTTGTACGTCCGTGGGCGACCTCCATGACTCAAGGC 1041
QY 322 TrpGlyHisArgSerAspThrGlnAlaAsnGluArgGlnCysProMetCysArgThrVal 341
DB 1042 TGGGGGTGCGGGGGGAGCGGGGCGCCAGAGCGGGAATGTCTCTGTCGCCCTGTGTG 1101
QY 342 GlyProTyValProLeuTrpLeuGlyCysGluAlaGlyPheTyrValAspAlaGlyPro 361
DB 1102 GGGCTTATGTGCTCTATGTGCTTGGCCAGAGCGCGGCTCTGCTGACCCCTGGGCGG 1161
QY 362 ProThrHisAlaPheThrProCysGlyYrHisValCysSerGlnLysSerAlaLysTyThr 381
DB 1162 CTTAGCCATCTTGTGACCTTGGCGCACGTGTCTGTGAGAACTGCGCCCTACTGG 1221
QY 382 SerGlnLysProLeuProHisGlyYrThrHisAlaPheHisAlaAlaCysProPheCysAla 401
DB 1222 GCCCAACACCACTTGGCCCGACCGCACCATGCTTTCATCCGCGCTTGGCGGG 1281
QY 402 ThrGlnLeuValGlyGluGlnAsnGlyLysLeuLysPheGlnGlyProIleAsp 420
DB 1282 GCCTGCTTACCGCGCAGCATGTGCTGCTCGCTCATTTTCCAGGCGCCGCTGTAG 1338

RESULT 2
US-10-750-185-40960/c
; Sequence 40960, Application US/10750185
; Publication No. US200502603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFIELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OR INVENTION: COMPOSITIONS FOR IMPROVING BOVINE TRAITS
; FILE REFERENCE: MM1100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 40960
; LENGTH: 1105
; TYPE: DNA
; ORGANISM: Bovine
US-10-750-185-40960

Alignment Scores:
Pred. No.: 4.41e-56 Length: 1105
Score: 711.50 Matches: 164
Percent Similarity: 66.9% Conservative: 48
Best Local Similarity: 51.7% Mismatches: 97
Query Match: 31.1% Indels: 12
DB: Gaps: 4

US-10-041-030-4 (1-420) x US-10-750-185-40960 (1-1105)

QY 108 ThrGlnSerProLysAspPheValValThrAspThrIleSerGlySerGlnAsnThrAsp 127
DB 1102 ACTGAAGTCTCATTAATTTTGTAAATACATGATATGTTCCAGAGCT-AAAAGTATGAT 1044
QY 128 -----GlnAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgLysIleValCys 145

```

```

Db      1043 AATTCGACACACAGCAGTACCAAGCACTATATCAAGATTCCTTTGAAATATATGT 984
      :
      :
      :
Qy      146 AapAaAgngLupProTyThraAlaArgIlePheAlaAlaGlyPheApsSerIyAaen 165
      :
      :
      :
Db      983 GAATATAAATTT-CCCTTTCATCATCAGATATATGTTGAGATTTGATCATCATCAAAAAAC 925
      :
      :
      :
Qy      166 IlPheLeuGlyGluLysAlaAlaLysTrpLysAsnProAspGlyHisMetAspGlyLeu 185
      :
      :
      :
Db      924 ATCTTTCTGGGGAAGAGAGTGGCATTTTGAAGGTGTCAAGATGGAGAGATGAAGGCTTGG 865
      :
      :
      :
Qy      186 ThrThraSngIValLeuValIleMetHisProArgGlyGlyPheThrGluGlnSerGlnPro 205
      :
      :
      :
Db      864 ACCACTAATGGCTGTTTGTATGATCCCATTAATGGGTTCCAGAAAGATTTTACACT 805
      :
      :
      :
Qy      206 GlyValTrpArgLulIeSerValCySgIAspValTyThr-LeuArgLulThraArgSe 225
      :
      :
      :
Db      804 GAATATGAGGGAATATATCAAGGTGTGAATGCACTGCTTGTATATATCAAAAT 745
      :
      :
      :
Qy      225 rAlaGlnGlnArgGlyLysLeuValGluSerGluThraAsnValLeuGlnAspGlySerIe 245
      :
      :
      :
Db      744 AGCTGACGAGAGAGAAATGGTGAAATGGAATCAGTAACCTTAAGCTGTCTTGT 685
      :
      :
      :
Qy      245 uIleAspLeuCySgIValIleThrLeuLeuTrpArgThraAlaAspGlyLeuPheHisThrPr 265
      :
      :
      :
Db      684 AATTCACCTGTCAGATTAACATCGTGTGGCATCTCCAGAAAGGCTTATTCG-CTTCC 626
      :
      :
      :
Qy      265 oThrgLulYshIsIleGluAlaLeuArgGlnGluIleAsnAlaAlaArgProGlnCySPr 285
      :
      :
      :
Db      625 TGTAGTGAATAAT---TCAGTTTAAAGAGAAATCAATTCAGTGTACCTCAGAACCC 569
      :
      :
      :
Qy      285 oValGlyLeuAsnThrLeuAlaPheProSerIleAsnAlaGlyGluValAlaGluLulY 305
      :
      :
      :
Db      568 TATCGGATTCAAAACACTGACATTCCTGTATGAAGAGAGAAAGAC---ATCATGA-AA 513
      :
      :
      :
Qy      305 sGlnProTrpAlaTyLleuSerCySgIYshIValIleGlyTyThrIleAsnTrpGlyHisAr 325
      :
      :
      :
Db      512 ACAGCCATGGGATATTTCTAACTGAGCCATGTGTGTGTATACCAATAT---GGAGAAA 456
      :
      :
      :
Qy      325 gSerAspThr-glualaAsnGluArgGluCyAsProMetCyAsArgThraValGlyProTyV 345
      :
      :
      :
Db      455 CAGAAAACATGATGAGAAAAGATTAAATGTCCTCAATGTATAGTCTGTGTCTTAC 396
      :
      :
      :
Qy      345 aLProLeuTrpLeuGlyCySgIAlaGlyPheTyrValAspAlaGlyProProThHisAr 365
      :
      :
      :
Db      395 TCTTCTGTGACTGTGATTTGAAGCTAGATTTGTGTGATCTGTGCTCCCAACTTACG 336
      :
      :
      :
Qy      365 laPheThrProCySgIYshIValIleSerGluLysSerAlaLysTyTrpSer-GlnIle 384
      :
      :
      :
Db      335 TATGTAGCACCTGTGTGTTAGGAGATTCAAGAAAATTAATCTCCAAATGGCCCAAGATC 276
      :
      :
      :
Qy      385 ProLeuProHisIleGlyThraIleAlaPheHisAlaAlaCyAsProPheCyAlaThrgLulE 404
      :
      :
      :
Db      275 CCACCTTTCATGCTGTCTATGTGTGTATGACAGCCCTGTGTGTGTGTGTGTGTG 216
      :
      :
      :
Qy      405 ValGlyGluGlnAsnCySgIleLysLeuIlePheGlnGlyProIle 419
      :
      :
      :
Db      215 GTTGAGAGAACAGGCTACATTACATTAATGCTCAAGAACCTTTA 171
      :
      :
      :

```

RESULT 3

```

US-10-750-623-40960/c
; Sequence 40960, Application US/10750623
; Publication No. US20050287531A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENFIELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM1100-1

```

```

; CURRENT APPLICATION NUMBER: US/10/750,623
; CURRENT FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; PRIOR FILING DATE: 2002-12-31
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 40960
; LENGTH: 1105
; TYPE: DNA
; ORGANISM: Bovine
; US-10-750-623-40960
; ORGANISM: Bovine 19866880722813

Alignment Scores:
Pred. No.: 4,41e-56 Length: 1105
Score: 711.50 Matches: 164
Percent Similarity: 66.9% Conservative: 48
Best Local Similarity: 51.7% Mismatches: 97
Query Match: 31.1% Indels: 12
DB: 8 Gaps: 4

US-10-041-030-4 (1-420) x US-10-750-623-40960 (1-1105)

Qy      108 ThrGluSerProIleAspPheValIleThraAspThrIleSerGlySerGlnAsnThrAsp 127
      :
      :
      :
Db      1102 ACTGAAAGCTCATTAATTTGTATTAATGATATAGTCCAGAGT-AAAGTATGAT 1044
      :
      :
      :
Qy      128 -----GluAlaGlnIleThrGlnSerThrIleSerArgPheAlaCyAsArgIleValCyS 145
      :
      :
      :
Db      1043 AATTCGACACACAGCAGTACCAAGCACTATATCAAGATTCCTTTGAAATATATGT 984
      :
      :
      :
Qy      146 AapAaAgngLupProTyThraAlaArgIlePheAlaAlaGlyPheApsSerIyAaen 165
      :
      :
      :
Db      983 GAATATAAATTT-CCCTTTCATCATCAGATATATGTTGAGATTTGATCATCATCAAAAAAC 925
      :
      :
      :
Qy      166 IlPheLeuGlyGluLysAlaAlaLysTrpLysAsnProAspGlyHisMetAspGlyLeu 185
      :
      :
      :
Db      924 ATCTTTCTGGGGAAGAGAGTGGCATTTTGAAGGTGTCAAGATGGAGAGATGAAGGCTTGG 865
      :
      :
      :
Qy      186 ThrThraSngIValLeuValIleMetHisProArgGlyGlyPheThrGluGlnSerGlnPro 205
      :
      :
      :
Db      864 ACCACTAATGGCTGTTTGTATGATCCCATTAATGGGTTCCAGAAAGCTTACACT 805
      :
      :
      :
Qy      206 GlyValTrpArgLulIeSerValCySgIAspValTyThr-LeuArgLulThraArgSe 225
      :
      :
      :
Db      804 GAATATGAGGGAATATATCAAGGTGTGAATGCACTGCTTGTATATATCAAAAT 745
      :
      :
      :
Qy      225 rAlaGlnGlnArgGlyLysLeuValGluSerGluThraAsnValLeuGlnAspGlySerIe 245
      :
      :
      :
Db      744 AGCTGACGAGAGAGAAATGGTGAAATGGAATCAGTAACCTTAAGCTGTCTTGT 685
      :
      :
      :
Qy      245 uIleAspLeuCySgIValIleThrLeuLeuTrpArgThraAlaAspGlyLeuPheHisThrPr 265
      :
      :
      :
Db      684 AATTCACCTGTCAGATTAACATCGTGTGGCATCTCCAGAAAGGCTTATTCG-CTTCC 626
      :
      :
      :
Qy      265 oThrgLulYshIsIleGluAlaLeuArgGlnGluIleAsnAlaAlaArgProGlnCySPr 285
      :
      :
      :
Db      625 TGTAGTGAATAAT---TCAGTTTAAAGAGAAATCAATTCAGTGTATCCTCAGAACCC 569
      :
      :
      :
Qy      285 oValGlyLeuAsnThrLeuAlaPheProSerIleAsnAlaGlyGluValAlaGluLulY 305
      :
      :
      :
Db      568 TATCGGATTCAAAACACTGACATTCCTGTATGAAGAGAGAAAGAC---ATCATGA-AA 513
      :
      :
      :
Qy      305 sGlnProTrpAlaTyLleuSerCySgIYshIValIleGlyTyThrIleAsnTrpGlyHisAr 325
      :
      :
      :
Db      512 ACAGCCATGGGATATTTCTAACTGAGCCATGTGTGTGTGTGTGTGTGTGTGTGTG 456
      :
      :
      :
Qy      325 gSerAspThr-glualaAsnGluArgGluCyAsProMetCyAsArgThraValGlyProTyV 345
      :
      :
      :
Db      455 CAGAAAACATGATGAGAAAAGATTAAATGTCCTCAATGTATAGTCTGTGTCTTAC 396
      :
      :
      :
Qy      345 aLProLeuTrpLeuGlyCySgIAlaGlyPheTyrValAspAlaGlyProProThHisAr 365
      :
      :
      :
Db      395 TCTTCTGTGACTGTGATTTGAAGCTAGATTTGTGTGATCTGTGCTCCCAACTTACG 336
      :
      :
      :

```

```
QY 365 lapherProCyGlyYHieValCyseRgluYseRialaLyTyTrpSer-GlnIle 384
Db 335 TAGTAGACACCTGTGTAGCGAAGTTCAAGAAAATAATCTCCATAGGACCCCAATC 276
QY 385 ProLeuProHiasglYThrhIaIaPheHiaIaIaCyAProPheCyAalArhGlnLeu 404
Db 275 CCACTTCTCATGTGTGTATGATGATGACCTGAGCCCTTGTGTGTGTCATGAGTTG 216
QY 405 ValGIGluGlnIaenCyseIleYleuIlePheGlnIleProIle 419
Db 215 GTTGAGACACAGGCTACATTAGCATTTATGCTCAAGACCTTTA 171

RESULT 4
US-09-925-065A-679103
/ Sequence 679103, Application US/09925065A
/ Publication No. US20040181048A1
/ GENERAL INFORMATION:
/ APPLICANT: Wang, David G.
/ TITLE OF INVENTION: Identification and Mapping of Single
/ FILE REFERENCE: 108827.135
/ CURRENT APPLICATION NUMBER: US/09/925, 065A
/ PRIOR FILING DATE: 2001-08-08
/ PRIOR APPLICATION NUMBER: US 60/243, 096
/ PRIOR FILING DATE: 2000-10-24
/ PRIOR APPLICATION NUMBER: US 60/252, 147
/ PRIOR FILING DATE: 2000-11-20
/ PRIOR APPLICATION NUMBER: US 60/250, 092
/ PRIOR FILING DATE: 2000-11-30
/ PRIOR APPLICATION NUMBER: US 60/261, 766
/ PRIOR FILING DATE: 2001-01-16
/ PRIOR APPLICATION NUMBER: US 60/269, 846
/ PRIOR FILING DATE: 2001-05-09
/ NUMBER OF SEQ ID NOS: 957086
/ SOFTWARE: FastSeq for Windows Version 4.0
/ SEQ ID NO 679103
/ LENGTH: 1345
/ TYPE: DNA
/ ORGANISM: Homo sapiens
US-09-925-065A-679103

Alignment Scores:
Pred. No.: 2,86e-17 Length: 1345
Score: 293.00 Matches: 80
Percent Similarity: 49.2% Conservative: 17
Best Local Similarity: 40.6% Mismatches: 50
Query Match: 12.8% Indels: 53
Gaps: 6
DB: 6

US-10-041-030-4 (1-420) x US-09-925-065A-679103 (1-1345)
QY 170 GlnUyeAlaAlaLySTpLySaenProAProGlyHieAeAProGlyLeuThrhAengly 189
Db 354 GACCGAGCGGCGCAATGCGGACCCGACATGCGCTGATGATGAGACTGACCAATGGA 413
QY 190 ValLeuValMetHisProAArgGlyGlyPheThrhGlnIuSerGlnProGlyValITPArg 209
Db 414 GTCTGTGTATGACCCGCGAGCGGCTCTCCAGGACTCAGCCCGGCTGTGCGCG 473
QY 210 GlnIleSerValCyseGlyAeAProValTyTrhLeuArgGlnThrhAeSerAlaGlnIaArg 229
Db 474 GAATCTCGGCTGTGTGGAATGTGTACATTTGCGGAGACGCGCTCAGCCGACGACGG 533
QY 230 GlnUyeLeuValGluSerGluThrhAenValLeuGlnAePlySerLeuIleAePlyCys 249
Db 534 GCGAAGCTGTGA----- 545
QY 250 GlnAlaThrhLeuThrpArgThrhAlaAePlyLeuPheHieThrhProThrhGlnUyeHie 269
Db 546 -----GCTGGCCCGCTCCATTC-CCCACC----- 568
QY 270 IleGlnAlaLeuArgGlnIuIleAenAlaAlaArgProGlnCyseProValGly----- 287
```

```
Db 569 ---CCTATCTTCCAGGCGCTCAAGAGTGGCCCATCCCC-TCGCCAAGCCAGAGATG 624
QY 288 LeuAenThrhLeuAlaPheProSerIleAenAglYgluValValGlnUyeGlnPro 307
Db 625 ATCTCCAGTCTTCTTCCCGGAGGAGCCCGCAAGAGCTCCATTCAGGAGATCCAGGCC 684
QY 308 TrpAlaTyTrpLeuSerCyseGlyHieValHieGlyTyThrhAenThrpGlnIleAeSerAe 327
Db 685 CAGGCGGCTTGTAGCATGACCACTTG-----TCACAGACAGAGC- 725
QY 328 ThrGlnAlaAenGlnUyeGluCysePrometCyseArgThrhValGlyPro----- 343
Db 726 AGACAGACAGGCTGCCCAAGCATAGCATATGACATGAGGTCCAGGAGGTGATGCG 785
QY 344 -----TyValProLeuThrpLeu---GlyCys 351
Db 786 ATCAAGGACACACAGACATCTGTGCACTGCCAGGTGCTTCAGAGCTGC 836

RESULT 5
US-10-750-185-38950
/ Sequence 38950, Application US/10750185
/ Publication No. US20050260603A1
/ GENERAL INFORMATION:
/ APPLICANT: MMI GENOMICS, INC.
/ APPLICANT: DENISE, Sue K.
/ APPLICANT: KERR, Richard
/ APPLICANT: ROSENFIELD, David
/ APPLICANT: HOLM, Tom
/ APPLICANT: BATES, Stephen
/ APPLICANT: FANTIN, Dennis
/ TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
/ FILE REFERENCE: MM1100-2
/ CURRENT APPLICATION NUMBER: US/10/750, 185
/ PRIOR FILING DATE: 2003-12-31
/ PRIOR APPLICATION NUMBER: US 60/437, 482
/ PRIOR FILING DATE: 2002-12-31
/ NUMBER OF SEQ ID NOS: 64922
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 38950
/ LENGTH: 3078
/ TYPE: DNA
/ ORGANISM: Bos taurus
US-10-750-185-38950

Alignment Scores:
Pred. No.: 34.4 Length: 3078
Score: 102.00 Matches: 88
Percent Similarity: 34.3% Conservative: 52
Best Local Similarity: 21.6% Mismatches: 160
Query Match: 4.5% Indels: 108
Gaps: 20
DB: 20

US-10-041-030-4 (1-420) x US-10-750-185-38950 (1-3078)
QY 13 AenUyeGluProValYerTyGlyGluLeuValIleGlnIleTyThrhAenglyAlaLeuPro 32
Db 1471 AACAGAGAG-----GAGCGCTCATTTGTTCCAAAGCAGGAAATTCATC 1515
QY 33 AenGlyAeP---ArgGlyAeArgYleSer-----ArgPheAlaLeuTyTrpYls 47
Db 1516 TTGAGAGACTCAAGAAACATTTGAGAGAGAGCCCTTGTGTGTGATTAATATTTCTGCT 1575
QY 48 ArgProValAeAenGlnUyeValYerProSerThrhValHieValIleSerThrhProGlnAla 67
Db 1576 AGACTCAAGCTCTCCAGTTTAAAGAAAGAACTTCAGCTGATTCACCTGTGTGCAT 1635
QY 68 SerUyeAlaIleSerCyseGlyGlnIleSerIleSerTyThrhLeuSerAeAeAenGln 87
Db 1636 GAAAGACCTTAAAGGCTTAAGAGAAAGTTG-----GCTGTGTGTATGAAGAT 1686
QY 88 ThrValValValGlnUyeThrhIleAePlyAePlyThrhAePhePhe---GlnValGlyArg 106
```



```

Db 1687 GCAGTCCCTAAGGTTGTTTCATTAATGTTCCCACTGATCTTAAGCAACTGTGAAA 1746
Qy 107 SerThrgluserProileasphevalValThraPthrIleSerGlySerGlnaThr 126
Db 1747 TCAACTTCCTTTCTTTTGATTTT-----ACAACCTGCTGAGATGTAACAAGTATCT 1800
Qy 127 Asp-----GluAlaGlnIleThrgInserThrlleSerArgPhe 139
Db 1801 GAACCTTATCAAGTTTTCATAAATGAAAAAGATATCTACACAAATAATTAATAACT 1860
Qy 140 AlaCysArgIleValCysAspArgAsnGluProTyThrlaArgIlePhealaalagly 159
Db 1861 CCAATTCAG-----GAAAGAAATTCCTCATCTAACAAGACATCATTTCCCTAGT 1911
Qy 160 PheAspSerSerIys-----AsnIlePheLeuGlyGlnValaalaValyTTP 175
Db 1912 ACCAATCTTCAACAAAGACATGTTAACTTTTACTTAACATGATCTGTGCAA--- 1968
Qy 176 LysAsnProAspGlyHsmetAspGlyLeuThrlaAsnGlyVal---LeuValMetHis 194
Db 1969 AGAAATTTTGTAGG---AAAGACATTAACACTGAAAATGAGTACTTATTAACGAA 2025
Qy 195 ProArgGlyGlyPheThrgInGluSerGln-----ProGlyValTPArgGluIleSer 212
Db 2026 AACATCTCAAAACATTACTCGAGGGGTGAAAATGTTCCAGATGATGAGATGTAATAAGT 2085
Qy 213 ValCysGlyAspValTyThrlaArgGluThr-----ArgSerAla 226
Db 2086 ACTGTGTTTATGACCAAGTTTGTGAGATGTTTAAAGAAATATGCTTGAACCAAG 2145
Qy 227 GlnGlnaArgGlyLysLeuValGluSerGluThraAsnValleuGlnaAspGlySerIle 246
Db 2146 GAGAGACATTAACAATCACTCTTAAGAACTACTCTCATTTATGACCCCAAAATAGTT 2205
Qy 247 AspleuGlyAlaThrlaLeuThrgThraArgThrlaAspGlyLeuPheHis----- 263
Db 2206 -----GGTCTCAACCACTACTCTCACTGATGATGATCAACATTTATTCA 2256
Qy 264 ---ThrProThrgInlyHsIleGluAlaLeuArgGlnGluIleAsnAlaalaArgPro 282
Db 2257 GAGAGCCCAACAGAAACAAATAAGAAATGCTGCAAAATTTGATATCTAGAGAA 2316
Qy 283 GlnCysProValGlyLeuAsnThrlaAlaPheProSerIleAsnAlaGlyValVal 302
Db 2317 AACTTCAGTTCTGCTGAAATAT-----AGTCTCAATTCACCTGCAATTTG 2364
Qy 303 GlnGluLysGlnProTPalaTyLeuSerCysGlyHsIvalHsGlyTyHsIleAsnTP 322
Db 2365 AATTATTAAGTTCC----- 2379
Qy 323 GlyHsIleArgSerAspThrgInAlaAsnGluArgGluCysProMetCysArgThValGly 342
Db 2380 -----AGATCTATGACTTAC----- 2394
Qy 343 ProTyValProleuThrleuGlyCysGluAlaGlyPheTyValaAspAlaGlyProPro 362
Db 2395 -----CCACCA 2400
Qy 363 ThrHsIalPheThrProCysGlyHsIleValCysSerGlyLysSerAlaTyTyTTPSer 382
Db 2401 ACTTACACA-----AGTGAATCTTTCACAAATTTCTAGGT 2436
Qy 383 GlnIleProleuProHsIleGlyThr 390
Db 2437 AAATATCTTTAAAGAGGCACT 2460

```

RESULT 6
 US-10-750-623-38950
 ; Sequence 38950, Application US/10750623
 ; Publication No. US20050287531A1
 ; GENERAL INFORMATION:
 ; APPLICANT: MMI GENOMICS, INC.
 ; APPLICANT: Denise, Sue K.

```

; APPLICANT: KERR, Richard
; APPLICANT: ROSENFELD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: METHODS AND SYSTEMS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MW1100-1
; CURRENT APPLICATION NUMBER: US/10/750,623
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 38950
; LENGTH: 3078
; TYPE: DNA
; ORGANISM: Bovine 1986680741510
; US-10-750-623-38950

Alignment Scores:
Pred. No.: 34.4 Length: 3078
Score: 102.00 Matches: 88
Percent Similarity: 34.3% Conservative: 52
Best Local Similarity: 21.6% Mismatches: 160
Query Match: 4.5% Indels: 108
Gaps: 20

US-10-041-030-4 (1-420) x US-10-750-623-38950 (1-3078)
Qy 13 AsnLeuGluProValLysTyGlyGluLeuValleuGlyTyHsAsnGlyAlaLeuPro 32
Db 1471 AACAGAG-----GAGCGCTATTTCTTCAAGACAGGAATTTTCATC 1515
Qy 33 AsnGlyAsp---ArgGlyArgGlySer-----ArgPheAlaLeuTyLys 47
Db 1516 TTGAGAGACCAAGAACATTAAGAGAGAGCCCTTGCTGATGATTAATATTCGCT 1575
Qy 48 ArgProLysAlaAsnGlyValLysProSerThrlaHsIvalleSerThrProGlnAla 67
Db 1576 AGACCTCAAGCTCTCCCAAGTTTAAAGAACCTTCAAGCTGATTCACCTGTCTCAT 1635
Qy 68 SerTyAlaIleSerCysLysGlyGlnHsSerIleSerTyThrlaLeuSerArgAsnGln 87
Db 1636 GAAAGACTTTAAAGCTTAAGAGAAAGTTCC-----GCTGTGATGAAAGAT 1686
Qy 88 ThrValValValGlyTyThrlaAspLysAspThrlaAspMetPhe---GlnValGlyArg 106
Db 1687 GCAGTCCCTAAGGTTGTTTCATTAATGTTCCCACTGATCTTAAGCAACTGTGAAA 1746
Qy 107 SerThrgluserProileasphevalValThraPthrIleSerGlySerGlnaThr 126
Db 1747 TCAACTTCCTTTCTTTTGATTTT-----ACAACCTGCTGAGATGTAACAAGTATCT 1800
Qy 127 Asp-----GluAlaGlnIleThrgInserThrlleSerArgPhe 139
Db 1801 GAACCTTATCAAGTTTTCATAAATGAAAAAGATATCTACACAAATAATTAATAACT 1860
Qy 140 AlaCysArgIleValCysAspArgAsnGluProTyThrlaArgIlePhealaalagly 159
Db 1861 CCAATTCAG-----GAAAGAAATTCCTCATCTAACAAGACATCATTTCCCTAGT 1911
Qy 160 PheAspSerSerIys-----AsnIlePheLeuGlyGlnValaalaValyTTP 175
Db 1912 ACCAATCTTCAACAAAGACATGTTAACTTTTACTTAACATGATCTGTGCAA--- 1968
Qy 176 LysAsnProAspGlyHsmetAspGlyLeuThrlaAsnGlyVal---LeuValMetHis 194
Db 1969 AGAAATTTTGTAGG---AAAGACATTAACACTGAAAATGAGTACTTATTAACGAA 2025
Qy 195 ProArgGlyGlyPheThrgInGluSerGln-----ProGlyValTPArgGluIleSer 212
Db 2026 AACATCTCAAAACATTACTCGAGGGGTGAAAATGTTCCAGATGATGAGATGTAATAAGT 2085

```

| | | | |
|----|------|--|-------------|
| QY | 213 | ValGlyGlyAspValIleThrLeuArgGlyIleThr-----ArgSerAla | 226 |
| DB | 2086 | ACTGCTTTGATGCCACAGCTTTTGGAGATGTTATTAAGAAATATGCTTGAACCAACGG | 2145 |
| QY | 227 | GlInGlnArgGlyIleValLeuValGlySerGluThrSerValLeuGlnAspGlySerIleIle | 246 |
| DB | 2146 | GAGAGGATTAACAATCCAGTTCCTTAAGAACCTACTCTCATTTATGACCCCAAAATGTT | 2205 |
| QY | 247 | AspLeuGlyGlyAlaThrLeuLeuThrArgThrAlaArgGlyLeuPheHis----- | 263 |
| DB | 2206 | -----GGATCTCACCCACTCTACTACGCTGCTGATGCATCAACATTTATTTCCA | 2256 |
| QY | 264 | ---ThrProThrGlnIleHisIleGluIleAlaLeuArgGlnGluIleAsnAlaIleArgPro | 282 |
| DB | 2257 | GGAGCCCCAACGAAACAAACAACTAAGAAATAGCTGCAAAATTTGATTTCTAGAGAAAGA | 2316 |
| QY | 283 | GlnCysProValGlyIleuAsnThrLeuAlaPheProSerIleAsnArgGlyGluValVal | 302 |
| DB | 2317 | AACCTCAGTTCTGGCTGATAATAT-----ACTCCAAATTCACAGCAATTTGG | 2364 |
| QY | 303 | GluGluIleGlnProThrAlaTyrLeuSerCysGlyHisValHisGlyTyrHisAsnThr | 322 |
| DB | 2365 | AATTATAAGTTCCTC----- | 2379 |
| QY | 323 | GlyHisAspSerAspThrGlnAlaAsnGlnArgGluCysProMetCysArgThrValGly | 342 |
| DB | 2380 | -----AAATCTAGTACTTAC----- | 2394 |
| QY | 343 | ProTyrValProLeuThrLeuGlyCysGluAlaGlyPheTyrValAspAlaGlyProPro | 362 |
| DB | 2395 | ----- | CCACCA 2400 |
| QY | 363 | ThrHisAlaPheThrProCysGlyHisValCysSerGluIleSerAlaIleTyrTrpSer | 382 |
| DB | 2401 | ACTCCACACA-----AGTGAGAACTTTCACCAATTTCTAGT | 2436 |
| QY | 383 | GlnIleProLeuProHisGlyThr | 390 |
| DB | 2437 | AAATATCCTTAAAGAGGACCT | 2460 |

RESULT 7

US-11-086-12/c

Sequence 12, Application US/11121086

Publication NO. US20050266459A1

GENERAL INFORMATION:

APPLICANT: POULSEN, TIM S.

APPLICANT: NIELSEN, KIRSTEN V.

TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES

FILE REFERENCE: 09138, 6000-00000

CURRENT APPLICATION NUMBER: US/11/121,086

CURRENT FILING DATE: 2005-05-04

PRIOR APPLICATION NUMBER: 60/567,570

PRIOR FILING DATE: 2004-05-04

NUMBER OF SEQ ID NOS: 107

SOFTWARE: PatentIn version 3.3

SEQ ID NO 12

LENGTH: 119160

TYPE: DNA

ORGANISM: Homo sapiens

US-11-121-086-12

Alignment Scores:

Pred. No.: 2,47e+03 Length: 119160

Score: 101.50 Matches: 61

Percent Similarity: 36.9% Conservative: 22

Best Local Similarity: 27.1% Mismatches: 98

Query Match: 4.4% Indels: 45

DB: 12 Gaps: 13

US-10-041-030-4 (1-420) x US-11-121-086-12 (1-119160)

QY 199 PheThrGlnGluSerGlnProGlyValTrpArgGluIleSerValCysGlyAspValTyr 218

|||||:||||| ||| |||:|||||

| | | | |
|--|----------|---|--------|
| Db | 15419 | TTGAGTACCTCCGCCAGGAAGCTTTTAACATCTATGCGCTGACTACGTAACGAAGTCACC | 15360 |
| Qy | 219 | ThrluudrgluthrArgSerlaaglInlaArglylvleuValglu-----Ser | 235 |
| Db | 15359 | TCAAAAGGAGCTACGCAAGGGTGTCAATAGCAGGCGATGCTTCACTGACCTGAACCTTAACAG | 15300 |
| Qy | 236 | GluthrAsnValleuGlInlaArglyserLeuIleAerpleCyeglyValatrhleuLeutr | 255 |
| Db | 15299 | TCTACAGAGGTGCTCCAGCAGGGCAGA---GTAGAACAAATGTGGGCC----- | 15255 |
| Qy | 256 | ArgthrAlaAerglyleuPheIlethrProthIngluPheIleaglAlaLeuArgIn | 275 |
| Db | 15254 | -----TTGAGTCCAGCCAAACGAGGGTTCAAATCCAGCGTGGCA | 15216 |
| Qy | 276 | GlulIleAsnAlaAlaArgPro--GlnCyAProValGlyLeuAsnthrLeuAlaPhePro | 294 |
| Db | 15215 | CACTTCCGACCTGAGACGACCTGTAAATCATGCACTTAGCCTCTGAGGCTGTGTTTATG | 15156 |
| Qy | 295 | SerIleAsnArglyserGluValValleuGlulysglInProthralaIytleu---SerCyv | 313 |
| Db | 15155 | CATCTCATATGAGAGG-----GTGAAGAAGCTGTGGCTTTGGAGACTTTGGGCTCTGT | 15102 |
| Qy | 314 | ---GlyIleValIleaglYtrHis-----AsnProglyIleAerSer | 326 |
| Db | 15101 | ACCCCCACCTCTCCAGCTTCCACCCACCTTATCCCTTAATATTGGATTTGAAAC | 15042 |
| Qy | 327 | AerThrGlulAlaangluArglylucyAProMetCyvAerThr-----ValglyProYr | 344 |
| Db | 15041 | GATTCCT---TCTAAGAGAAAGAGAAAGCAGCAGAAATGGGAAATGACACAGCTCCAGA | 14985 |
| Qy | 345 | ValProleuThrleuGllyserGluAlaaglYrPheYrVal-----AspAlaGlyPro | 361 |
| Db | 14984 | GTGCTTACCTTGGACGAGCTGCTG--TATGAGATTCAAGTATCATGCTGAGCCCTGGCT | 14926 |
| Qy | 362 | ProthrIleAlaPheThrProcySglYrIleValCySerGluysSerAlaIyYrTrp | 381 |
| Db | 14925 | GGTCCCACTCAAGTCCT-----CACACACATTATTGG | 14893 |
| Qy | 382 | SerGlnIleProleuProHleaglYtrHisAlaPheHleAlaAlaCyvProPheCyAl | 401 |
| Db | 14892 | TTCGACATGCTCCAGCGCCCGCAGCGGACCCACCTG-----ACAGTGTCCCCCACTGCAG | 14839 |
| Qy | 401 | atrhInleuVal | 405 |
| Db | 14838 | GATATGCTCTGTA | 14826 |
| RESULT 8 | | | |
| US-11-121-086-14/c | | | |
| ; Sequence 14, Application US/11121086 | | | |
| ; Publication No. US20050266459A1 | | | |
| ; GENERAL INFORMATION: | | | |
| ; APPLICANT: POULSEN, TIM S. | | | |
| ; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES | | | |
| ; FILE REFERENCE: 09138.6000-00000 | | | |
| ; CURRENT APPLICATION NUMBER: US/11/121, 086 | | | |
| ; PRIOR FILING DATE: 2005-05-04 | | | |
| ; PRIOR APPLICATION NUMBER: 60/567, 570 | | | |
| ; NUMBER OF SEQ ID NOS: 107 | | | |
| ; SOFTWARE: PatentIn version 3.3 | | | |
| ; SEQ ID NO 14 | | | |
| ; LENGTH: 167891 | | | |
| ; TYPE: DNA | | | |
| ; ORGANISM: Homo sapiens | | | |
| US-11-121-086-14 | | | |
| Alignment Scores: | | | |
| Pred. No.: | 3.65e+03 | Length: | 167891 |
| Score: | 101.50 | Matches: | 61 |
| Percent Similarity: | 36.9% | Conservative: | 22 |
| Best Local Similarity: | 27.1% | Mismatches: | 98 |
| Query Match: | 4.4% | Indels: | 45 |

```

DB: 12 Gaps: 13
US-10-041-030-4 (1.4-20) x US-11-121-086-14 (1.1-167891)
Oy 199 PheThrgIngluSerGlnProGlyValTPArGluIleSerValCyseGlyAspValTyr 218
Db 166774 TTCAAGTACTCTCCACAGAACTGTTTAAACATCATATGCTGCTACGTAACAAGTCAC 166715
Oy 219 ThrLeuArGluThrArgSerAlaGlnIleArgGlyValLeuValGlu-----Ser 235
Db 166714 TCAAAAGGAACTAACCCAAAGGGGTGCTCAATGACGAGCAATGCTTCACTAACCAATCTAACAG 166555
Oy 236 GluThrAnuValLeuGlnAspGlySerLeuIleAspLeuCyGlyAlaThrLeuLeuTyr 255
Db 166654 TCTAACACAGGTCTCCACAGAGGACAGAA-----GTAAACAATATGTGGACC----- 166610
Oy 256 ArgThrAlaAspGlyLeuPheHisThrProThnGlnLysHisIleGluAlaLeuArgGln 275
Db 166609 -----TTTAGAGTCCAGCCAAACAGAGGTTCAAAATCCACAGCTGCCA 166571
Oy 276 GluIleAsnAlaAlaArgPro---GlnCyProValGlyLeuAsnThrLeuAlaPhePro 294
Db 166570 CACTTCGAGGCTGGAAGACCTGAAACATGCACTAGCTGCTGAGGCTCTGTTTAAAG 166511
Oy 295 SerIleAsnArgGlyLeuValValGluGluLysGlnProTPAlaTyrLeu---SerCy 313
Db 166510 CATCTCATAGAGGG-----GTGAAGACATGTGGCTTGGGAGACTTGGGCTCTCTGT 166457
Oy 314 ---GlyHisValHisGlyTyrHis-----AsnTrpGlyHisArgSer 326
Db 166456 ACCCCCACTCTCTCAGCTTCAACCACTTTATCCCTTAATATTTGGATTTCAGAAC 166397
Oy 327 AsPThrGluAlaAsnGlnArgGlyLysProMetCysArgThr-----ValGlyProTyr 344
Db 166396 GATCTC---TCTAAGAAAGAGAGAGACAGAGATGGAGAAAGTAAAGTCAACAAGCTCAAG 166340
Oy 345 ValProLeuThrPLeuGlyCyseGluAlaGlyPheTyrVal-----AspAlaGlyPro 361
Db 166339 GTGCTTACCTTGCAGAGGGCTGCTG-TATGATTCAGAGTATCATCTGAGCCCTGGGCT 166281
Oy 362 ProThrHisAlaPheThrProCyseGlyHisValCyseSerGluLysSerAlaLysTyrTrp 381
Db 166280 GGTCCCACTAAGTCTT-----CACACACTTAATTTGG 166248
Oy 382 SerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCyseProPheCysAl 401
Db 166247 TTCCCATGTCCCCACAGCCCAAGGGAGACCACTG-----ACAGTGTCCCCCACTGGCAG 166194
Oy 401 aThrGlnLeuVal 405
Db 166193 GATATGCTCTTGTA 166181

RESULT 9
US-11-013-247A-16
; Sequence 16, Application US/11013247A
; Publication No. US2005025496A1
; GENERAL INFORMATION:
; APPLICANT: GAO, ZEREN
; APPLICANT: SHEPPARD, PAUL O.
; APPLICANT: FOX, BRIAN A.
; APPLICANT: HOLLOWAY, JAMES T.
; APPLICANT: JASPERS, STEPHEN R.
; TITLE OF INVENTION: ZNF12, A TUMOR NECROSIS FACTOR
; FILE REFERENCE: 03-23
; CURRENT APPLICATION NUMBER: US/11/013,247A
; PRIOR FILING DATE: 2004-12-15
; PRIOR APPLICATION NUMBER: 60/530,125
; NUMBER OF SEQ ID NOS: 47
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 16
; LENGTH: 1664
; TYPE: DNA

```

| ORGANISM: Homo sapiens | FEATURE: | NAME/KEY: CDS | LOCATION: (28) ... (1617) | US-11-013-247A-16 |
|--|----------|--|---------------------------|-------------------|
| Alignment Scores: | | | | |
| Pred. No.: | 23.5 | Length: | 1664 | |
| Score: | 100.50 | Matches: | 102 | |
| Percent Similarity: | 34.8% | Conservative: | 52 | |
| Best Local Similarity: | 23.1% | Mismatches: | 170 | |
| Query Match: | 4.4% | Indels: | 119 | |
| DB: | 12 | Gaps: | 25 | |
| US-10-041-030-4 (1-420) x US-11-013-247A-16 (1-1664) | | | | |
| QY | 39 | ArglySerArgPheAlaLeuTyLysArgProLysAlaEnglyValLysProSerThr | 58 | |
| DB | 302 | AGAAATGAATTCAGGCCAGTGAATCCCCCA-----GTGTCCCAAGCA | 352 | |
| QY | 59 | ValHisValIleSerThrProGlnAlaSerLysAlaIle-----SerCysLysGly | 75 | |
| DB | 353 | AGGCTCAGAGATGCACATCCATGCATCTCGAACCTGTGCAGGTGAAGATGCCGAC | 412 | |
| QY | 76 | GlnHisSerLysEryr-----ThrLeuSerArgAsnGlnThrValValGlu | 92 | |
| DB | 413 | CCAGCGATCTCTCTCTCTCTCCGCCACCTTCTCCACATCAGTGAAGTCTTC | 472 | |
| QY | 93 | TyrThrHisAspLysAspThrAspMetPheGlnValGlyArgSerThrGluSerProIle | 112 | |
| DB | 473 | TTTCGT-----GTGTTCGTGCCGCCGCCGCCGCCGCCGCCGCTTGTTC | 514 | |
| QY | 113 | AspPheValValThrAspThrIleSerGlySerGlnAsnThrAspGlnAlaGlnIle--- | 131 | |
| DB | 515 | GCCGCGCTGTGCACACGACGAGCGCGCGGAGAACTACAGCAATCCCGCTACGC | 574 | |
| QY | 132 | ThrGlnSerThrIleSerArgPheAlaCysArgIleValCysAspArgAsnGluProTyr | 151 | |
| DB | 575 | ACTCAGTCTTCGTAACCCACATGCAGCCAAA-----AGATGTCACAACTAC | 622 | |
| QY | 152 | ThrAlaArgIlePheAlaIleGlyPheAspSerSerLysAsnIlePheLeuGlyGluLys | 171 | |
| DB | 623 | ACCGAGTGCCTTTC-----TTTGATCAG----- | 645 | |
| QY | 172 | AlaAlaLysTrpLysAsnProAspGlyHisMetAspGlyLeuThrThrAsnGlyValLeu | 191 | |
| DB | 646 | -----GAGAACCGGATTCCTACTGTGAG-----GAGAGAGACAACTGCC | 687 | |
| QY | 192 | ValMetHisProArg-----GlyGlyPheThrGluGlu----- | 202 | |
| DB | 688 | TTCCCGATCCCAAGTACCCACGTCGCGGTGCGGGGGTTTATACAGAGAGCGGGCTTG | 747 | |
| QY | 203 | ---SerGlnProGlyValTrp---ArgGlnIleSerValCysGly----- | 215 | |
| DB | 748 | CCCTCCCAATGGGAGCTGTGGGGCCACAGGGGTGATCTGGCCAGTCTGCCACCAACC | 807 | |
| QY | 216 | AspValTyrThrLeuArgGluThrArgSerAlaGlnArgGlyLysLeuValGlnSer | 235 | |
| DB | 808 | TCTCTCTACCTGTACCTGAGCTGCGCTGCATGCCCAAGCT-----GTAGAGGCC | 858 | |
| QY | 236 | GluThrAsnVal---LeuGlnAsp-----GlySerLeuIle-AspLeuCysGly | 250 | |
| DB | 859 | AGGCTGAGTGAAGGCTGCATCTTATGGGCGCACGGTATCCCATCCGAC---TGGGG | 915 | |
| QY | 250 | ValAlaThrLeuLeuThrArgThrAlaAspGly-----LeuPheHisThrProThrGlnLys | 268 | |
| DB | 916 | GAGCA-----ATGTGAGAGCTGACAGAGTGGCTCTGTCCACCACTCCCAACCGGCTG | 969 | |
| QY | 268 | HisIleGluAlaLeu----- | 273 | |
| DB | 970 | CCCCCTAACCCCTCTTGTGGTCCCGGTGGGAGACAGCCTTACCCCTCAGTGGGCTGATG | 1022 | |
| QY | 274 | -----ArgGlnGluIleAsnAlaIleAspPr | 282 | |

```

Db      1030 CTGATATACCTCTGGATTCAGCGGGCTCTGGACCGAGGGCTTTAGCGGGCTCTGGC 1089
Oy      282 oGlnCyAPro-----ValGlyLeuasnThrIleuAlaPh 293
Db      1090 TCGGGTGTCCCGAAGCGCGGGCTGAGGGCCAGGGCTGCGGGAGACCACTCTCCAGAG 1149
Oy      293 eProSerIeasnrglyuValValGluGluuValProThrAlaTyIleuSerCy 313
Db      1150 TCCCAACGACAGAGCGCTGTGGAGCGAGCTTATGGCCAGTCCACGCGAGTCTC----- 1201
Oy      313 sGlyHleValHleGlyTyHniValasnThrGlyHniArgSerAerThrGluAlaAsnGlyuAr 333
Db      1202 ----CCACCCATCCAGGAAAGCTTGG-----GCTACAGCTCC 1236
Oy      333 gGlnCyAProMetCyAArgThrValGlyProTyValProIleuThrIleuGlyCyAVal 353
Db      1237 CAGGACCCCGGTGAGGTGGCGGGCTGGCGGAGCTGAGCTAGGAGTGGCTC---TGCCGGCC 1293
Oy      353 aGlyPheTyValAlaSerAlaGlyProProThrHniValaPheThrProCyAelHniValCy 373
Db      1294 TGGGGTCTCTGACTACCTCTGGCTCTCCATCAGGTGTGGACGAGGCTCCCATCAAGG 1353
Oy      373 sSerGlnuSerAlaValuTyTrpSerGlnIleProIeuProHniGlyThrHniAlaPh 393
Db      1354 ACCCAAGCTCCAAAGTCAAGTGTGTGTCCTCCCACTCCCAAGCTTGGCCCAAGTCCAG 1413
Oy      393 eHniAlaAlaCyAProPhe-----CyAAlaThrGlnIleuValGlyGluHniAsy 410
Db      1414 GCTGGGAGCCCTGGCCCTCCCGCCGACATGTTGTGCCATCAGCGGAAATCCAGGGGGCC 1473
Oy      410 sIle 411
Db      1474 AATG 1477

RESULT 10
US-11-121-086-4/c
; Sequence 4, Application US/11121086
; Publication No. US20050266459A1
; GENERAL INFORMATION:
; APPLICANT: POULSEN, TIM S.
; APPLICANT: NIELSEN, KIRSTEN V.
; TITLE OF INVENTION: NUCLEIC ACID PROBES AND NUCLEIC ACID ANALOG PROBES
; FILE REFERENCE: 09138, 6000-00000
; CURRENT APPLICATION NUMBER: US/11/121,086
; CURRENT FILING DATE: 2005-05-04
; PRIOR APPLICATION NUMBER: 60/567,570
; PRIOR FILING DATE: 2004-05-04
; NUMBER OF SEQ ID NOS: 107
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 4
; LENGTH: 164810
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-121-086-4

Alignment Scores:
Pred. NO.:      8.39e+03      Length:      164810
Score:          97.50      Matches:      68
Percent Similarity: 33.8%      Conservative: 21
Best Local Similarity: 25.9%      Mismatches: 89
Query Match:      4.3%      Indels:      85
                        Gaps:      14

US-10-041-030-4 (1-420) x US-11-121-086-4 (1-164810)

Oy      175 TrpIys-AsnProAspGly-HisMetAspGlyLeuThrThrAsnGlyValLeuValMetH 194
Db      49348 TGGAAAGCTCCAAAGACGCGCCACAACTGGGCTTAACTAAGAGGGC-----CCAGGCG 49295
Oy      194 isPProArgGlyGlyPheThrGluGluSerGlnProGlyValITPrAArgGlnIleSerValC 214
Db      49394 ACTCGAGGGGAGAC-----GCCAGGTGGCAAGATGCCAAAGGCT 49256

```

```

QY      214  ysglyAarValYrYrThleuAagGluThArXserAlaGlnGlnAryGlyuVleuValG 234
Db      49255  GC-----CGGCTCTCAAGCCGCCCGGAGGAGC-----49227
QY      234  luserGluThrAsnValleuGlnAarGlySerIleuIleApleuCyGlyAlaThrLeuL 254
Db      49226  -----ACCCGATTCAGCTCAAGCT-----49206
QY      254  eutrP-----AgtThrAlaAarGlyleuPheHisThrProThrGlnlyVhIleGluA 272
Db      49205  --TGGCAATGCAACCCCTTCAGAGGTTTCGTCTCAAGACTCGCGTGGCAACGTCGCG 49148
QY      272  lalauAagGlnGluIleAsnAlaAlaArgPro-----GlnCyPro-Val 286
Db      49147  AGGTC---CAAGCCTCACTGTCAAGCTTCTCCGCTACAAATGATGCCAGGCCCGAGAG 49091
QY      287  glyLeuAanThrleuAlaPheProSerIleAsnArglyAgluValGluGluVwGln 306
Db      49090  GGCACTTAAGCCGTCGGCCACCCCGCTAAGAGAGCTCGGGGTGCTG-----49043
QY      307  ProThrAlaTyrlleuSerCyS-----Gly 314
Db      49042  CCCGGAGCTGGCTGTCTACGCCGCACTCCGCGGGCTCAAGACACTGTGAACGCGGC 48983
QY      315  HlValHlAaglyTyrlHlAenTPrglyHlAaXserAarP-ThrGluAlaAaGluAarG 334
Db      48992  TCTGTGATGACACAGCAAGAGTGGAGAGCGCGACGACGACGACCAAGCGCGCGCG 48923
QY      334  uCyAarPmetCyAarGThrValAlaTyProTyValAlaProleuTrpleuGlyCyAgluAa 354
Db      48922  GCCAAGTCCGAAAGGGGAGACATTCCTCGTGGGCACTTGACATCAAGCTTACCTTGCGG 48863
QY      354  yPhe-----TyrlAarAlaGlyProProThrHlAaPheThrProCy 369
Db      48862  CCATATCTTGTGTGCCACGTGGCGGGGCTGGCTGTGCCACGAT-----48817
QY      369  sGlyHlAaValCyAserGluuYserAlaTyTyTrTPserGlnIleProleuProHlAa 389
Db      48816  -----TGGACGCTTGGGACAAACCGGCTCCCACTGG 48782
QY      389  yThrHis-----AlaPheHlAaAlaCyAProPheCyAalAthrGlnleuVa 405
Db      48781  CTGGCACTGGAGCTGTCCCGGCGCTGGCTGTCTGTCTCC-----GAGGCTCAAGCTGCT 48728
QY      405  1 405
Db      48727  T 48727
RESULT 11
US-10-330-773-362/C
; Sequence 362, Application US/10330773
; Publication No. US2006040262A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001300
; CURRENT APPLICATION NUMBER: US/10/330, 773
; CURRENT FILING DATE: 2002-12-27
; NUMBER OF SEQ ID NOS: 981
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 362
; LENGTH: 21728
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: m1ec feature
; LOCATION: (1) ... (21728)
; OTHER INFORMATION: n = A,T,C or G
US-10-330-773-362
Alignment Scores:

```


QY 238 AsnValleuGlnAerGlySerLeuLeuAerLeuCyseGlyValAthrLeuLeuTrpArgThr 257
Db 43644 -----TGGGGGGCT----- 43652
QY 258 AlaAerGlyLeuPheHisThr---ProThrGlnLysHisIleGluAla----- 272
Db 43653 GCAGGGCTCTGTGTAAAGAGAGCCGCTCAGAGTCCCAACCCACAGCATTCGAGGGGA 43712
QY 273 -----LeuArgGlnGluIleAsnAlaAlaArgProGlnCysProValGly 287
Db 43713 GTCTCTCTGTAAATGAAGAGCCGCTCAGAGTCCCAACCCACAGCATTCGAGGGGA 43772
QY 288 LeuAsnThrLeuAlaPheProSerIleAsnArgLysGluValGluGlnLysGlnPro 307
Db 43773 AACCATGGAAGCATGATGGAAGACAGACTGATGGCTCAGCTCCGAGGAAAGTCCA 43832
QY 308 -----TrpAlaTrp----- 310
Db 43833 GAGCGCTCTCATTAAGCAGCGCCGGTTGGTCACTCAATGCAATGAAGTCACTCT 43892
QY 311 -----LeuSerCyseGlyHisValHisGlyThrHisAsnTrp-GlyHisArgse 326
Db 43893 CTCTGCGCCACTGTGCTGTAAACCCACGCTGGGAGGACATTCCTCCGATTCAGT 43947
QY 326 rAerThrGluAlaAsnGlnArgLysCysPrometCysArgThrValGlyProThrValPr 346
Db 43948 -----CCTTTTCCCCACAGGGAGACACAGGACTTAATCTCA 43985
QY 346 oleuTrpLeuGlyCyseGlnAlaGlyPheThrValAspAlaGlyProThrHisAlaPr 366
Db 43986 TTGGCTGCTGGGGTCCAGGCC-----TACCCCTGACATCAGC 44024
QY 366 eThrPro---CyseGlyHisValCysSerGlnLysSerAlaLysTrpSerGlnIlePr 385
Db 44025 AACAGCCAGGAGGCCACCGT-----TGGTCCCTTAACCC 44060
QY 385 oleuProHis-----GlyThrHis 391
Db 44061 GCTAACCCCACTGGGGTAATAAGGGTCACT 44094
RESULT 13
US-11-112-908-55
; Sequence 55, Application US/11112908
; Publication No. US20050260659A1
; GENERAL INFORMATION:
; APPLICANT: Harrie, Cole
; TITLE OF INVENTION: Breast Cancer Biomarkers
; FILE REFERENCE: 04-164-US
; CURRENT APPLICATION NUMBER: US/11/112,908
; PRIOR FILING DATE: 2005-04-22
; PRIOR APPLICATION NUMBER: US 60/564,758
; PRIOR FILING DATE: 2004-04-23
; PRIOR APPLICATION NUMBER: US 60/575,978
; PRIOR FILING DATE: 2004-06-01
; PRIOR APPLICATION NUMBER: US 60/631,702
; PRIOR FILING DATE: 2004-11-30
; PRIOR APPLICATION NUMBER: US 60/633,826
; PRIOR FILING DATE: 2004-12-07
; NUMBER OF SEQ ID NOS: 511
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 55
; LENGTH: 193789
; TYPE: DNA
; ORGANISM: Homo sapiens
US-11-112-908-55

DB: 12 Gape: 12
US-10-041-030-4 (1-420) x US-11-112-908-55 (1-193789)
QY 178 ProAerGlyHisMetAerGlyLeuThrThrAerGlyValLeuValMetHisProArgGly 197
Db 93685 CCAGAGGGCCACAGAGTGGGCCGAGATCTGGGCTGCAGAGCAATTCACAGCGTGA 93744
QY 198 GlyPheThrGlnLysSerGlnProGlyValTrpArgGluIleSerValCysGlyAerVal 217
Db 93745 GGTGCACAAATGAGGACAGATCTGGGGCTCCGTGAGACTCTCGATGAGGT----- 93798
QY 218 TyrThrLeuArgGlnThrArgSerAlaGlnArgGlyLysLeuValGluSerGluThr 237
Db 93799 -----AGAGGGGCGCTGGGACAGGGCTGGGGC----- 93828
QY 238 AsnValleuGlnAerGlySerLeuLeuAerLeuCyseGlyValAthrLeuLeuTrpArgThr 257
Db 93829 -----TGGGGGGCT----- 93837
QY 258 AlaAerGlyLeuPheHisThr---ProThrGlnLysHisIleGluAla----- 272
Db 93838 GCAGGGCTCTGTGTAAACCCACGCTGGGAGGACATTCCTCCGATTCAGT 93897
QY 273 -----LeuArgGlnGluIleAsnAlaAlaArgProGlnCysProValGly 287
Db 93898 GTCTCTCTGTAAATGAAGAGCCGCTCAGAGTCCCAACCCACAGCATTCGAGGGGA 93957
QY 288 LeuAsnThrLeuAlaPheProSerIleAsnArgLysGluValGluGlnLysGlnPro 307
Db 93958 AACCATGGAAGCATGATGGAAGACAGAACTGATGGCTCAGCTCCGAGGAAAGTCCA 94017
QY 308 -----TrpAlaTrp----- 310
Db 94018 GAGCGCTCTCATTAAGCAGCGCCGGTTGGTCACTCAATGCAATGAAGTCACTCT 94077
QY 311 -----LeuSerCyseGlyHisValHisGlyThrHisAsnTrp-GlyHisArgse 326
Db 94078 CTCTGCGCCACTGTGCTGTAAACCCACGCTGGGAGGACATTCCTCCGATTCAGT 94132
QY 326 rAerThrGluAlaAsnGlnArgLysCysPrometCysArgThrValGlyProThrValPr 346
Db 94133 -----CCTTTTCCCCACAGGGAGACACAGGACTTAATCTCA 94170
QY 346 oleuTrpLeuGlyCyseGlnAlaGlyPheThrValAspAlaGlyProThrHisAlaPr 366
Db 94171 TTGGCTGCTGGGGTCCAGGCC-----TACCCCTGACATCAGC 94209
QY 366 eThrPro---CyseGlyHisValCysSerGlnLysSerAlaLysTrpSerGlnIlePr 385
Db 94210 AACAGCCAGGAGGCCACCGT-----TGGTCCCTTAACCC 94245
QY 385 oleuProHis-----GlyThrHis 391
Db 94246 GCTAACCCCACTGGGGTAATAAGGGTCACT 94279
RESULT 14
US-11-183-136-41
; Sequence 41, Application US/11183136
; Publication No. US20060019896A1
; GENERAL INFORMATION:
; APPLICANT: Li, Dean
; TITLE OF INVENTION: NETRIN-RELATED COMPOSITIONS AND USES
; FILE REFERENCE: UTH-P01-011
; CURRENT APPLICATION NUMBER: US/11/183,136
; PRIOR FILING DATE: 2005-07-14
; PRIOR APPLICATION NUMBER: US 60/587,796
; PRIOR FILING DATE: 2004-07-14
; NUMBER OF SEQ ID NOS: 44
; SOFTWARE: FASTSEQ for Windows Version 4.0
; SEQ ID NO 41
; LENGTH: 1770

Alignment Scores:
Pred. No.: 1.12e+04 Length: 193789
Score: 97.00 Matches: 61
Percent Similarity: 30.2% Conservative: 15
Best Local Similarity: 24.2% Mismatches: 85
Query Match: 4.2% Indels: 91

```

TYPE: DNA
ORGANISM: Mouse
US-11-183-136-41

Alignment Scores:
Pred. No.: 59 Length: 1770
Score: 96.50 Matches: 92
Percent Similarity: 33.2% Conservative: 38
Best Local Similarity: 23.5% Mismatches: 140
Query Match: 4.2% Indels: 122
DB: 12 Gaps: 21

US-10-041-030-4 (1-420) x US-11-183-136-41 (1-1770)

QY 36 ArgGlyArgGlySerArgPheAlaLeuTyrIleArgProIleAlaAsnGlyValIle 55
DB 291 CAGGAGAGATGAGGAGCTGCTACTCTACTGCAAGGCT-----329
QY 56 ProSerThrValIleSerThrProGlnAlaSerIleValIleSerCysIleGly 75
DB 330 -----CACGTGAGTGGCTTACCCAG-----TCCACTAGAGGC 362
QY 76 GlnHis-SerIleSerThrThrLeuSerArgAsnGlnThr-----ValValValGlnTyr 93
DB 363 CACATACCCCTCTCATGAAACAAGACGTGAGTTACAGACACTGCTGTGACTTT 422
QY 93 rThrHisAspIleAspThrAspMet-----PheGlnValIleIleArgSerTh 108
DB 423 TGAGTATGCGCGCCACGGCTCATGCTCTGAGAAAGTCCCTGACAAATGGCCGC--AC 479
QY 108 rGluSerProIleAspPheValValThrAspThrIleSer-----GlySerGlnAsnTh 126
DB 480 CTGGCAGCCCTTACAGTCTATGACAGAGACTGATGACAGGCTTGGCATGTCTGCCG 539
QY 126 rAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgIleValCysAs 146
DB 540 ACCTGCCCGAGACATGACCTTCCAGCCGCCACCGGCTGCTGC-----565
QY 146 rArgAsnGlnProTyrThrAlaArgIlePheAlaAlaGlyPheAspSerSerIleAsnI 166
DB 586 -----ACCGAGAGATACTCAGCC-----TGGGCGAGGCTCCAAAGAAAGACATGT 632
QY 166 e-----PheLeuGlyGlyIleValAlaIleValIleValIleValIleValIleVal 183
DB 633 GCGCTTGAAGTAAAGGAGCGCTTGGCATTTTGGCCGCTTACCTGCGTGAATGGA 692
QY 183 rGlyLeuThrThrAsnGlyValLeuValIleMetHisProArgGlyValIlePheThrGlnGln 203
DB 693 CAACCTGTACACAG-----ATGGAGAG 716
QY 203 rGlnProGlyValIlePheArgGlnIleSerValCysGlyAspValIleThrLeuArgGlnTh 223
DB 717 GCGCAAGGGCTC-----AAGAGATTCTTCACTTCACTGACTGACT 755
QY 223 rArgSerAlaGlnGlnIleArg-----GlyIleValLeuValGlnSerGlnThr----- 237
DB 756 GCGCATGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 815
QY 238 -----AsnValLeuGlnAspGlySerIleIleAspLeuCys-GlyAlaThrLeu 254
DB 816 CAAGTACTTTCATGCAATCTCCATATTCAGAGTATTCGCAAGGTGAGTGAAGTCAACTGCA 875
QY 254 eutTrpArgThrAlaAspGlyLeuPheHisThrProThrGlnIleValHisIleGlyValAlaLeu 274
DB 876 TG-----CCAACCTGTGCACTGCGAG-----898
QY 274 rGlnGlnIleAsnAlaAlaArgProGlnCysProValIleGlyLeuAsnThrLeuAlaPhe 294
DB 899 -----AGGGCAGCCCTGACGTGAGTGTGAACA-----928
QY 294 roSerIleAsnArgIleValValGlnGlnIleValGlnIleValGlnIleValGlnIleValGln 314
DB 929 -----ACACCAAGGCGCCCACTGTGCGAG 953

```

```

QY 314 IyHisValHisGlyTyrHisAsnTrpGlyHisArgSerAspThrGlnAlaAsnGlnIleArg 334
DB 954 GT-----GCAGAGAAACTTCCGCAACAGCGCTGGCGAGC 989
QY 334 IuCySPrometCysArg-----ThrValGlyProTyrValPro- 346
DB 990 TGGCTCTTACTGCTCCCTGCGCCACAGGCTCTCCCATATGCTGTGGCGCGGCTCCGCG 1049
QY 347 LeuTrpLeuGlyCysGlnAlaGlyPheTyrValAlaAspAlaGlyProProThrHisAlaPhe 366
DB 1050 CTTGGCAGTCAAGACCAAGCAAGCACTACTAT-----GGCCCCCTTGGGAGACGCTC 1100
QY 367 ThrProCysGlyHisValCysSerGluIleSerAlaIleTyrThrSerGlnIleProLeu 386
DB 1101 CTTCGCGCCCAAGTGTCTCCAGTGCAGAGACTGTAGCTATCTGTGCTGTGCTCTTC 1160
QY 387 Pro---HisGlyThrHisAlaPheHisAla 395
DB 1161 CCAAGCCAGACTCTACGCTTTTCAGACT 1190

RESULT 15
US-10-750-185-32846
; Sequence 32846, Application US/10750185
; Publication No. US200502603A1
; GENERAL INFORMATION:
; APPLICANT: MMI GENOMICS, INC.
; APPLICANT: DENISE, Sue K.
; APPLICANT: KERR, Richard
; APPLICANT: ROSENBLUD, David
; APPLICANT: HOLM, Tom
; APPLICANT: BATES, Stephen
; APPLICANT: FANTIN, Dennis
; TITLE OF INVENTION: COMPOSITIONS FOR INFERRING BOVINE TRAITS
; FILE REFERENCE: MM11100-2
; CURRENT APPLICATION NUMBER: US/10/750,185
; PRIOR FILING DATE: 2003-12-31
; PRIOR APPLICATION NUMBER: US 60/437,482
; NUMBER OF SEQ ID NOS: 64922
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 32846
; LENGTH: 1363
; TYPE: DNA
; ORGANISM: Bovine 1986680713026
US-10-750-185-32846

Alignment Scores:
Pred. No.: 67.1 Length: 1363
Score: 94.50 Matches: 58
Percent Similarity: 35.0% Conservative: 28
Best Local Similarity: 23.6% Mismatches: 78
Query Match: 4.1% Indels: 83
DB: 8 Gaps: 13

US-10-041-030-4 (1-420) x US-10-750-185-32846 (1-1363)

QY 205 ProGlyValIlePheArgGlu-----210
DB 27 CCGGCGCTTGAAGAGAGAGCTGGGTCTCTGAGATGTCCTCCAGAGAGCTTTGGGAGAC 86
QY 211 -----IleSerValCysGlyAspValIleThrLeuArgGlnIleThrArgSerAlaGln 227
DB 87 CACGTCTCTTCCAGACTTGGGTGTGCTGTGGGCTGAGAGGTGAGCGGCTCCCGCC 146
QY 228 GlnArgGlyIleValLeuValGlnSerGlnIleThrAsnValLeuGlnAspGlySerLeuIleAsp 247
DB 147 TCCAGACCAACCTTCAAGGCGCCCGCTGAG--CTCTTGTGCAAGGCGGTCTCCGCGCA 204
QY 248 IeuCysGlyAlaThrIleLeuThrArgThrAlaAspGlyLeuPheHisThrProThrGln 267
DB 205 TTGTGC-----210

```


| | | | |
|----|-----|---|-----|
| Oy | 268 | lyshhsliegualaleumarglingluileanlaialaaryproglincysvprovalgly | 287 |
| | | | |
| Db | 211 | ---CACACCAGGACCAATATGTCCTCCATCAAGTCTGCCGACGAAATCTGCCAGTAGGT | 267 |
| | | | |
| Oy | 288 | leuamnthleualapheproserileamarglyval--- | 305 |
| | | | |
| Db | 268 | GTT-----GTTAATGTACCATTTTACAGAGAAACCACTAGAGGCTCAGAGAAAGCA | 321 |
| | | | |
| Oy | 306 | -----GlnProTyr--AlaTyrleusercysvalylvalinhis | 317 |
| | | | |
| Db | 322 | ATGATGTCGTCGAGGGCTCGCCGCGAGGTGGAGGACAGAGCTGAATCTCA-----AAC | 372 |
| | | | |
| Oy | 318 | glytyrhisantprgly---hisargseraerpthglualaleamgluarglycysvpro | 336 |
| | | | |
| Db | 373 | TCTGATCACTCTGGGCCCCGACCTCAAGTCACAGACAGCTCTTCCAGAGGCCACTTC | 432 |
| | | | |
| Oy | 337 | MetCysvargthrValglYProTyrValProleuttrpleuglycysgluialaglyPhe--- | 355 |
| | | | |
| Db | 433 | AGCTGCTAGAAATCTCAAACTTAC-----TGTAGGAGAGGTGTGTGTC | 474 |
| | | | |
| Oy | 356 | -----TyrValaIaerAlaglyProPro-----Thr | 363 |
| | | | |
| Db | 475 | AGAGCCCTGTGTCTTTTCCCATCTATCTCTCTGTCAACACCTGAGACAGAGCCTTTGCAGAAg | 534 |
| | | | |
| Oy | 364 | hisalaperthrProCysglYhisValCysSerGlu--- | 375 |
| | | | |
| Db | 535 | CACAGAGAGAGCCCTGCGCGTGTGGGTGTCTTGAAGGCTCTGTCTCGCGCGTCTCTC | 594 |
| | | | |
| Oy | 376 | ---LysSerAlaIaYserTyrtrpserGlnileProleuProhiValglYthrhis---Ala | 392 |
| | | | |
| Db | 595 | CTCTCTGGCGGGTCCCCAGAAACAGACAGCAAGCTCCCGGGCGGTGTGGTATCCCATAGACC | 654 |
| | | | |
| Oy | 393 | PhehiValaIaCysvpro | 398 |
| | | | |
| Db | 655 | CAGCACAAGGCCAGGCTT | 672 |

Search completed: March 2, 2006, 07:44:06
Job time : 1418 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceeleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: March 2, 2006, 06:06:38 ; Search time 1066 Seconds
(without alignments)
3258.104 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MESPGEHCAPKPEPVXG.....ATQVGEQNCIKLIFQGPID 420

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 9793542 seqs, 4134689005 residues
Total number of hits satisfying chosen parameters: 19587084

Minimum DB seq length: 0
Maximum DB seq length: 2000000000
Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-O=abs/ABSSWB.spool/US10041030/rnuc 01032006 134411 21967/app_query.fasta_1
-DB=Published Applications NA Main -OPMT=fastap -SUFFIX=p2n.rnpsm
-NMATCH=0.1 -LOOPT=0 -LOOPTXT=0 -UNITS=bits -START=1 -END=1
-MATRIX=blomsun62 -TRANS=human40.cdi -LIST=45 -DOCCALIGN=200 -THR_SCORE=pct
-THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL -OUTFMT=pct -NORM=ext
-HEAPSIZE=500 -MAXLEN=2000000000 -HOST=abss05h
-USER=US10041030 @CGN 1.1 1026 @runat 01032006 134411 21967 -NCPU=6 -ICPU=3
-NO MAP -NEG SCORES=0 -WAIT -DSBLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MAP TIMEOUT=30 -THREDS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database: Published Applications NA Main:*

- 1: /cgn2_6/prodata/1/pubpna/US07_PUBCOMB.seq:*
- 2: /cgn2_6/prodata/1/pubpna/US08_PUBCOMB.seq:*
- 3: /cgn2_6/prodata/1/pubpna/US09_PUBCOMB.seq:*
- 4: /cgn2_6/prodata/1/pubpna/US09B_PUBCOMB.seq:*
- 5: /cgn2_6/prodata/1/pubpna/US10A_PUBCOMB.seq:*
- 6: /cgn2_6/prodata/1/pubpna/US10B_PUBCOMB.seq:*
- 7: /cgn2_6/prodata/1/pubpna/US10C_PUBCOMB.seq:*
- 8: /cgn2_6/prodata/1/pubpna/US10D_PUBCOMB.seq:*
- 9: /cgn2_6/prodata/1/pubpna/US10E_PUBCOMB.seq:*
- 10: /cgn2_6/prodata/1/pubpna/US11_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|-------|-------------|--------|-------|--------------------------------------|
| 1 | 2290 | 100.0 | 1263 | 6 | US-10-085-117-288 Sequence 288, App |
| 2 | 2290 | 100.0 | 5579 | 6 | US-10-197-666A-135 Sequence 135, App |
| 3 | 2290 | 100.0 | 5579 | 6 | US-10-085-117-287 Sequence 287, App |
| 4 | 2290 | 100.0 | 5579 | 6 | US-10-041-030-4 Sequence 3, App1 |
| 5 | 2283 | 99.7 | 1263 | 3 | US-09-843-905A-7 Sequence 7, App1 |
| 6 | 2283 | 99.7 | 1263 | 6 | US-10-317-250-7 Sequence 7, App1 |
| 7 | 2283 | 99.7 | 1263 | 7 | US-10-258-703-7 Sequence 7, App1 |

| | | | | | | |
|----|--------|------|--------|----|---------------------|--------------------|
| 8 | 2175.5 | 95.0 | 1260 | 3 | US-09-843-905A-5 | Sequence 5, App1 |
| 9 | 2175.5 | 95.0 | 1260 | 6 | US-10-317-250-5 | Sequence 5, App1 |
| 10 | 2175.5 | 95.0 | 1260 | 6 | US-10-085-117-285 | Sequence 285, App |
| 11 | 2175.5 | 95.0 | 1260 | 7 | US-10-258-703-5 | Sequence 5, App1 |
| 12 | 2111.5 | 92.2 | 1717 | 5 | US-10-197-666A-133 | Sequence 133, App |
| 13 | 2111.5 | 92.2 | 1734 | 6 | US-10-085-117-284 | Sequence 284, App |
| 14 | 1917 | 83.7 | 1257 | 3 | US-09-843-905A-3 | Sequence 3, App1 |
| 15 | 1917 | 83.7 | 1257 | 6 | US-10-317-250-3 | Sequence 3, App1 |
| 16 | 1917 | 83.7 | 1257 | 7 | US-10-258-703-3 | Sequence 3, App1 |
| 17 | 1917 | 83.7 | 1257 | 7 | US-10-041-030-1 | Sequence 1, App1 |
| 18 | 1917 | 83.7 | 1257 | 6 | US-10-443-108-9 | Sequence 9, App1 |
| 19 | 1912 | 83.5 | 1257 | 3 | US-09-843-905A-1 | Sequence 1, App1 |
| 20 | 1912 | 83.5 | 1257 | 6 | US-10-317-250-1 | Sequence 1, App1 |
| 21 | 1912 | 83.5 | 1257 | 7 | US-10-258-703-1 | Sequence 1, App1 |
| 22 | 1642 | 71.7 | 1338 | 6 | US-09-843-905A-11 | Sequence 11, App1 |
| 23 | 1642 | 71.7 | 1338 | 6 | US-10-317-250-11 | Sequence 11, App1 |
| 24 | 1642 | 71.7 | 1338 | 6 | US-10-258-703-11 | Sequence 11, App1 |
| 25 | 1482.5 | 64.7 | 2508 | 6 | US-10-104-047-1958 | Sequence 1958, App |
| 26 | 1302 | 56.9 | 2966 | 10 | US-11-097-143-10850 | Sequence 10850, A |
| 27 | 1075 | 46.9 | 723 | 9 | US-10-450-763-4485 | Sequence 4485, App |
| 28 | 1064 | 46.5 | 202802 | 6 | US-10-085-117-286 | Sequence 286, App |
| 29 | 1048 | 45.8 | 155579 | 6 | US-10-085-117-283 | Sequence 283, App |
| 30 | 1022 | 44.6 | 541 | 6 | US-10-029-386-22982 | Sequence 22982, A |
| 31 | 996 | 42.5 | 31882 | 10 | US-11-097-143-10849 | Sequence 10849, A |
| 32 | 974 | 42.5 | 528 | 6 | US-10-029-386-20504 | Sequence 20504, A |
| 33 | 808 | 35.3 | 24173 | 3 | US-09-764-891-9284 | Sequence 9284, App |
| 34 | 808 | 35.3 | 24173 | 5 | US-10-091-572-855 | Sequence 855, App |
| 35 | 775 | 33.8 | 465 | 3 | US-09-867-550-1947 | Sequence 1947, App |
| 36 | 763 | 33.3 | 514 | 6 | US-10-029-386-6790 | Sequence 6790, App |
| 37 | 763 | 33.3 | 514 | 6 | US-10-029-386-9281 | Sequence 9281, App |
| 38 | 717 | 31.3 | 468 | 3 | US-09-918-995-15544 | Sequence 15544, A |
| 39 | 678 | 29.6 | 377 | 3 | US-09-867-550-1719 | Sequence 1719, App |
| 40 | 516 | 22.5 | 520 | 9 | US-10-756-149-249 | Sequence 249, App |
| 41 | 516 | 22.5 | 812 | 3 | US-09-764-891-7046 | Sequence 7046, App |
| 42 | 516 | 22.5 | 813 | 3 | US-09-764-891-7047 | Sequence 7047, App |
| 43 | 516 | 22.5 | 813 | 3 | US-09-764-891-7047 | Sequence 7047, App |
| 44 | 343 | 15.0 | 1587 | 9 | US-10-450-763-4484 | Sequence 4484, App |
| 45 | 330.5 | 14.4 | 6014 | 9 | US-10-450-763-13370 | Sequence 13370, A |

ALIGNMENTS

RESULT 1
US-10-085-117-288
; Sequence 288, Application US/10085117
; Publication No. US2003023334A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCES: 529452004, 21
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 288
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-085-117-288

Alignment Scores:

Pred. No.: 8.2e-255
Score: 2290.00
Percent Similarity: 100.0%
Best Local Similarity: 100.0%
Query Match: 100.0%
DB: 6
US-10-041-030-4 (1-420) x US-10-085-117-288 (1-1263)

Qy 1 MetPheSerProGlyGlnGluGlnHleCYaAlaProAsnLysGluProValLysTyrGly 20
Db 1 ATGTTTCCCTGGCCGAGGAAACATGGGCCCCCAATAAGAGCCAGTGAATAATACGGG 60
Qy 21 GlnLeuValValLeuGlyTyrAaenGlyAlaLeuProAsnGlyAaPaTgGlyAaGAGLys 40
Db 61 GACCTGTGTGGTCCGGGTGCAATAGTGTCTTACCAATGAGATAGAGACGAGAGAAA 120
Qy 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
Db 121 AGTAGATTGGCTCTTACAGCGGCCCCAGGCAATAGTGTCAAAACCCAGCACCTGTCAT 180
Qy 61 ValLleSerThrProGlnAlaSerLysAlaLleSerCYaLysGlyGlnHisSerLleSer 80
Db 181 GTGATATCCAGGCCCGAGGATCCAGAGCTATACGCTGCAAAAGGTTCACACAGATATCC 240
Qy 81 TyrThrLleuSerArgaenGlnThrValValValGluTyrThrHisAaPlyAaPThrAaP 100
Db 241 TACACTTTGTCAAGAAATCAGACTGTGTGTGAGTGAACACACATGATTAAGCATACGGAT 300
Qy 101 MetPheGlnValGlyArgSerThrGlyLysProLleAaPheValValThrAaPThrLle 120
Db 301 ATGTTTCAAGTGGGCAATCAACAGAAAGCCCTATTCGACTTGTTCACAGACACATTC 360
Qy 121 SerGlySerGlnAaenThrAaPGLuaGlnLleThrGlnSerThrLleSerArgPheAla 140
Db 361 TCTGGCAGCCAGAAACAGAGAGAGCCAGATCCACAGAGACCAATATCCAGGTTGCC 420
Qy 141 CYaArgLleValCYaAaPaTgAaenGluProTyrThrAlaArgLlePheAlaAaGlyPhe 160
Db 421 TGCAGAGTCTGTGTGCGACAGAAATGAACCTTACACACAGAGATATTCGCGCGGATTT 480
Qy 161 AaPSeSerLysAaenLlePheLysGlyGluValAlaLysTyrLysAaPProAaPGLy 180
Db 481 GACCTTCCAAACATATTTCTTGAAGAAAGGCAACAAAGTGAAGAAACCCCGAGGCG 540
Qy 181 HisAaPGLyLeuThrThrAaenGlyValLeuValAaenHisAaPProAaPGLyPheThr 200
Db 541 CACATGATGGGCTCACTAATGAGGCTCTGTGTATGATCCACAGGGGCTTACCC 600
Qy 201 GlnGluSerGlnProGlyValTTPaGGLuLleSerValCYaGlyAaPValTyrThrLleu 220
Db 601 GAGGAGTCCAGCGCCGGGCTGTGGCGGAGATCTCTGTGTGAGATGTGTACACCTTG 660
Qy 221 ArgGluThrArgSerAlaGlnGlnAaArgLysLeuValGlnSerGluThrAaenValLeu 240
Db 661 CGAGAAACAGGTGCGCCAGCAAGAGAAAGCTGTGAGAAATGAGACCAAGCTCTG 720
Qy 241 GlnAaPGLySerLleuLleAaPLeuCYaGlyAlaThrLleuLeuTTPaGGLuAaPGLy 260
Db 721 CAGGACGGCTCCCTCACTTGACTGTGTGTGGGCCACTCTCTCTGTGAGAAACAGCATGG 780
Qy 261 LeuPheHisThrProThrGlnLysHisLleGlnAlaLeuArgGlnLleAaenAla 280
Db 781 CTTTTCATATCTCCAACTCAGAAAGCACATAGAACCTCCGCGAGAGATTAAGCCGCC 840
Qy 281 ArgProGlnCYaProValGlyLeuAaenThrLleuAlaAaPProSerLleAaenGlyLeu 300
Db 841 CGGCTCAGTGTCTGTGGGCTCAACACCTGGCTTCCCGAGCATCAACAGAAAGAG 900
Qy 301 ValValGlnGluLysGlnProTTPaLleTyrLleuSerCYaGlyHisValHisGlyTyrHis 320
Db 901 GTGGTGAAGAGAGAGAGCCCTGGGCAATATCTCACTTGTGGCCACGTGCACGGGTACAC 960
Qy 321 AaenTTPaLysHisAaPSeAaPThrGlnAlaAaenGluArgLysCYaProMetCYaArgThr 340
Db 961 AACTGGGGCCATCGAGTGAACAGAGGCCAAGAGAGAGGTGTCCCATGTGAGACT 1020
Qy 341 ValGlyProTyrValProLeuThrLysGlyCYaGlyValAaGlyPheTyrValAaPGLy 360
Db 1021 GTGGGCCCCATAGTCTCTGTGGCTGTGTGAGAGAGATTTTATGAGCGCAGGA 1080
Qy 361 ProProThrHisAlaPheThrProCYaGlyHisValCYaSerGluLysSerAlaLysTyr 380

Db 1081 CCGCAACTCATGCTTCACTCCCTGTGACAGTGTGCTCGGAGAAAGTCTGCAAAATAC 1140
Qy 381 TTPSerGlnLleProLeuProHisGlyThrHisAlaPheHisAlaAlaCYaProPheCYa 400
Db 1141 TGGTTCAGATCCCGTGTCTCATGGAACCTCAAGCATTTCAACGCTGTGCCCTTCTGT 1200
Qy 401 AlaThrGlnLeuValGlyGlnGlnAaenCYaLleLysLeuLlePheGlnGlyProLleAaP 420
Db 1201 GCTACACAGCTGTGGGAGAGAAACCTGATCAATTAATTTTCCAGGTCCAAATTGAC 1260
RESULT 2
US-10-197-666A-135
; Sequence 135, Application US/1019766A
; Publication No. US20030092037A1
; GENERAL INFORMATION:
; APPLICANT: ASahi KASEI KABUSIKI KAISYA
; TITLE OF INVENTION: ELK1 phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197,666A
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,104
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 135
; LENGTH: 5579
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (177) .. (1436)
US-10-197-666A-135
Alignment Scores:
Pred. No.: 6.8e-254 Length: 5579
Score: 2290.00 Matches: 420
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
Gaps: 0
US-10-041-030-4 (1-420) x US-10-197-666A-135 (1-5579)
Qy 1 MetPheSerProGlyGlnGluGlnHleCYaAlaProAsnLysGluProValLysTyrGly 20
Db 177 ATGTTTCCCTGGCCGAGGAAACATGGGCCCCCAATAAGAGCCAGTGAATAATACGGG 226
Qy 21 GlnLeuValValLeuGlyTyrAaenGlyAlaLeuProAsnGlyAaPaTgGlyAaGAGLys 40
Db 237 GACCTGTGTGGTCCGGGTGCAATAGTGTCTTACCAATGAGATAGAGACGAGAGAAA 296
Qy 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
Db 297 AGTAGATTGGCTCTTACAGCGGCCCCAGGCAATAGTGTCAAAACCCAGCACCTGTCAT 356
Qy 61 ValLleSerThrProGlnAlaSerLysAlaLleSerCYaLysGlyGlnHisSerLleSer 80
Db 357 GTGATATCCAGGCCCGAGGATCCAGAGCTATACGCTGCAAAAGGTTCACACAGATATCC 416
Qy 81 TyrThrLleuSerArgaenGlnThrValValValGluTyrThrHisAaPlyAaPThrAaP 100
Db 417 TACACTTTGTCAAGAAATCAGACTGTGTGTGAGTGAACACACATGATTAAGCATACGGAT 476

QY 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValValThrAspThrIle 120
 DB 477 ATGTTTCAGGTTGGCGCATCAACAGAAAGCCCTATTCACATTCCTGTCACAGACACATTT 536
 QY 121 SerGlySerGlnAsnThrAspGluValGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 537 TCTGGCAGCCAGAAACGAGACGAAAGCCCATATCAACAGACCATATTCAGGTTGCC 536
 QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160
 DB 597 TGCAGATGCTGTGGCAGAGAAATGAACCTTACACAGACGAGATATTCGCCGCCGAGATT 656
 QY 161 AspSerSerIleAsnIlePheLeuGlyGluValAlaAlaIleTyrPheAsnProAspGly 180
 DB 657 GACTCTTCCAAAACATATTTCTTGAGAAAGGACCAAGTGGAAAAACCCCGACGGC 716
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleValIleMetHisProArgGlyGlyPheThr 200
 DB 717 CACATGGATGGGCTCACTACTAATGCGCTCTGTGATGCAATCCAGAGGGGGCTTACC 776
 QY 201 GluGlnSerGlnProGlyValITrParGluIleSerValCysGlyAspValTyrThrLeu 220
 DB 777 GAGGAGTCCAGACCCGGGCTGGCGGAGATCTCTGTGTGAGATGTGTACACTTG 836
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyValLeuValGluSerGluThrAsnValLeu 240
 DB 837 CGAGAAACCGAGTGGCGCCAGCAACGAGAAAGCTGTGAAAGTGACCAACGCTCTG 896
 QY 241 GluAspGlySerLeuIleAspLeuCysGlyValIleThrLeuThrProArgThrAlaAspGly 260
 DB 897 CAGGACCGCTCTCATTAACCTGTGTGGGCGCATCTCTCTGGAACACGAGAGGG 956
 QY 261 LeuPheHisThrProThrGlnIleValHisIleGluValLeuArgGlnIleAsnAlaAla 280
 DB 957 CTTTTCATCTCCAACTCAGAGACCATGAGAGCCCTCCGCGAGAGATTACGCCGC 1016
 QY 281 ArgProGlnCysProValIleGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLeu 300
 DB 1017 CGGCTCAGTGTCTGTGGGGCTCAACACCTGCTCTCCCGACATCAACAGAAAGAG 1076
 QY 301 ValValGluGluValGlnProTrpAlaTyrLeuSerCysGlyValHisValIleGlyThrHis 320
 DB 1077 GTGTGGAGGAGAAAGCCCTGGGCAATCTCACTGTGTGCGCACGTCGCGGTACAC 1136
 QY 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGluValGluCysProMetCysArgThr 340
 DB 1137 AACTGGGGCCATGGAGTGAACGAGAGCCACAGAGGAGAGTGTCCCATGTGCAGACT 1196
 QY 341 ValGlyProTyrValProLeuTrpLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
 DB 1197 GTGGGGCCCTATGCTCTCTGTGGCTGTGGCTGTGAGCAGAGATTATTTATGAGCGCAGGA 1256
 QY 361 ProProThrHisAlaPheThrProCysGlyValIleValCysSerGlyValSerAlaValTyr 380
 DB 1257 CCGCACAATGATGCTTCACTCTCTGTGACACGCTGTGCGAGAGAGTGTGCAAAATAC 1316
 QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaIlePheHisAlaIleCysProPheCys 400
 DB 1317 TGGTCTCAATCCCGTTGGCTCATGAACTCAATGATTTCAAGCTGTGCTTCTGTCTGT 1376
 QY 401 AlaThrGlnLeuValGlyGlnGlnAsnCysIleValLeuIlePheGlnGlyProIleAsp 420
 DB 1377 GCTAACAGCTGGTGGGAGCAAACTGCATCAATTAATTTTCCAGGTCCCAATTGAC 1436

RESULT 3
 US-10-085-117-287
 ; Sequence 287, Application US/10085117
 ; Publication No. US20030232334A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Morris, David W.
 ; APPLICANT: Engelhard, Eric K.
 ; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER

FILE REFERENCE: 529452000121
 ; CURRENT APPLICATION NUMBER: US/10/085,117
 ; CURRENT FILING DATE: 2002-02-27
 ; PRIOR APPLICATION NUMBER: US 09/798,586
 ; PRIOR FILING DATE: 2001-03-02
 ; NUMBER OF SEQ ID NOS: 361
 ; SOFTWARE: FastSeq for Windows Version 4.0
 ; SEQ ID NO: 287
 ; LENGTH: 5597
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 ; US-10-085-117-287
 Alignment Scores:
 Pred. No.: 6,83e-254 Length: 5597
 Score: 2290.00 Matches: 420
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-041-030-4 (1-420) x US-10-085-117-287 (1-5597)
 QY 1 MetPheSerProGlyGlnGlnGluHisCysAlaProAsnIleGluProValIleTyrGly 20
 DB 177 ATGTTTCCCTGGCGAGAGACCTGGCCCCCAATAGAGAGCCAGTGAATACGGG 236
 QY 21 GluLeuValIleGluGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgArgLeu 40
 DB 237 GACTGTGTGTGTGGGTGCAATGTGTCTTACCAATGAGATGAGAGAGAGAGAA 256
 QY 41 SerArgPheAlaLeuTyrIleAspProValIleAsnGlyValIleProSerThrValHis 60
 DB 297 AGTAGATTGGCTCTTACAAAGCGGCCCAAGGCAAAAGTGTCAAAACCGACACCGTCCAT 356
 QY 61 ValIleSerThrProGlnAlaSerIleValIleSerCysValGlyGlnHisSerIleSer 80
 DB 357 GTATATTCACGCGCCAGAGATCAAGGCTATCACTGAGTCAAGGTCAACAGTATATCC 416
 QY 81 TyrThrLeuSerArgAsnGlnThrValIleValIleGluTyrThrHisAspIleAspThrAsp 100
 DB 417 TACACTTGTCAAGATCAAGATCACTGTGTGTGATACACATCATATAGATACGAT 476
 QY 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValIleThrAspThrIle 120
 DB 477 ATGTTTCAGGTGGGAGATCAACAGAAAGCCCTATGACTGTGTGTCAACAGACTT 536
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 537 TCTGGCAGCCAGAAACGAGACGAAAGCCCATATCAACAGACCATATTCAGGTTGCC 586
 QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160
 DB 597 TGCAGATGCTGTGGCAGAGAAATGAACCTTACACAGACGAGATATTCGCCGCCGAGATT 656
 QY 161 AspSerSerIleAsnIlePheLeuGlyGluValAlaAlaIleTyrPheAsnProAspGly 180
 DB 657 GACTCTTCCAAAACATATTTCTTGAGAAAGGCGCAGAAAGTGGAAAAACCCCGACGGC 716
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleValIleMetHisProArgGlyGlyPheThr 200
 DB 717 CACATGATGGGCTCACTACTAATGCGCTCTGTGATGCAATCCAGAGGGGGCTTACC 776
 QY 201 GluGlnSerGlnProGlyValITrParGluIleSerValCysGlyAspValTyrThrLeu 220
 DB 777 GAGGAGTCCAGACCCGGGCTGGCGGAGATCTCTGTGTGAGATGTGTACACTTG 836
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyValLeuValGluSerGluThrAsnValLeu 240
 DB 837 CGAGAAACCGAGTGGCGCCAGCAACGAGAAAGCTGTGAAAGTGACCAACGCTCTG 896
 QY 241 GluAspGlySerLeuIleAspLeuCysGlyValIleThrLeuThrProArgThrAlaAspGly 260

Dh 897 CAGACGGCTCCCTCATTTGACCTGTGTGGGCGCACTCTCTCTGGAGAACAGCATGGG 956
Qy LeuphenisthrProthrnglnlyshlelleglnaleuargnglnuileanala 280
Dh 261 CTTTTCATCTCCACTCGAACACATGAAACCTTCGGGAGGAGATTTAACGGCCG 1016
Qy 957 CTTTTCATCTCCACTCGAACACATGAAACCTTCGGGAGGAGATTTAACGGCCG 1016
Dh 281 ArgProGlnCyseProValGlyLeuasnThrleuAlaPheProserlleaamrglysglu 300
Qy 1017 CGGCGCTCAGTGTCTGTGGGGCTCAACACCTGGCTTCCCGACATCAACAGAAAGAG 1076
Dh 301 ValValGlnGlnlysglnProtrPalaTytleuSerCyseGlyshValshlellyThrs 320
Qy 1077 GTGTGTGAGAGAGAACGAGCCCTGGGCAATCTCACTGTGTGGCCACGTCGACGGGTAC 1136
Dh 321 AsnTrpGlyshArgSerArpThrGlnAlaAsnGlnuArgGluCyseProMetCyseArgThr 340
Qy 1137 AACGGGGCCATCGAGATGACACGAGGCCAACAGAGGAGTGTCCCATGTGCAGACT 1196
Dh 341 ValGlyProTyValProleuTrpLeuGlyCyseGlnuAlaGlyPheTyValAspAlaGly 360
Qy 1197 GTGGGCGCCATATGTGCTCTGTGGCTTGGCTGTGAGCAGGATTTTATGACGACGGA 1256
Dh 361 ProProThrshAlaPheThrProCyseGlyshValCyseSerGlnlySerAlaIshTy 380
Qy 1257 CCGGCACTCATGTCTTCACTCCCTGTGACACGTCGTGCGAGAGTCTGCAAAATAC 1316
Dh 381 TrpSerGlnleProleuProHleGlyThrsAlaPhehAlaAlaCyseProPheCys 400
Qy 1317 TGGTTCAGATCCCGTGTGCTCATGAGATCATCATTCACGGTGTGCTTGTCTGT 1376
Dh 401 AlaThrGlnleuValGlyGlnGlnAsnCyseIlelyshleuIlePheGlnGlyProIleAsp 420
Qy 1377 GCTACACAGCTGGTGGGGAGCAAACTGCATCAATTAATTTCCAGGTCGATTTGAC 1436

RESULT 4

US-10-041-030-3
; Sequence 3, Application US/10041030
; Publication No. US2002150934A1
; GENERAL INFORMATION:
; APPLICANT: Powers, Scott
; APPLICANT: Mu, David
; APPLICANT: Xiang, Phil
; APPLICANT: Peng, Yue
; APPLICANT: Talarik Inc.
; TITLE OF INVENTION: Diagnosis and Treatment of Cancer Using Mammalian
; FILE REFERENCE: 018761-00681005
; CURRENT APPLICATION NUMBER: US/10/041,030
; PRIOR FILING DATE: 2001-12-28
; PRIOR APPLICATION NUMBER: US 60/259,502
; PRIOR FILING DATE: 2001-01-02
; NUMBER OF SEQ ID NOS: 42
; SOFTWARE: Patent Ver. 2.1
; SEQ ID NO 3
; LENGTH: 5921
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: CDS
; LOCATION: (501)..(1763)
; OTHER INFORMATION: human pellino 2
US-10-041-030-3

Alignment Scores:

Pred. No.: 7.4e-254 Length: 5921
Score: 2290.00 Matches: 420
Percent Similarity: 100.0% Conservative: 0
Best Local Similarity: 100.0% Mismatches: 0
Query Match: 100.0% Indels: 0
DB: 5 Gaps: 0

US-10-041-030-4 (1-420) x US-10-041-030-3 (1-5921)

Qy 1 MetPheSerProGlyGlnGlnGlnuIshCyseValaProAsnlysgluProVallyTyrgly 20
Dh 501 ATGTTTCCCTGCGCCAGAGGAAACATGCGGCCCAATTAAGGACGAGTAAATACGGG 560
Qy 21 GlnleuValValleuGlyTyraenglyAlaLeuProAsnGlyshArgGlyTyraengly 40
Dh 561 GACCTGTGTGTCTGGGTACATGTGTCTTACCAATGGAGATGAGAGACGAGAA 620
Qy 41 SerArgPheAlaLeuTyrlshArgProLyshAlaAsnGlyVallyshProserThrs 60
Dh 621 AGTAAATTTGCCCTTCAACAGCGGCCCAAGCAATGTGTCAAAACCCAGACCTTCAT 660
Qy 61 ValIleSerThrProGlnAlaSerlyshAlaIleSerCyseGlyshGlnIshSer 80
Dh 681 GTATATCCACGCGCCAGGATCCAAAGCTATCAGCTGCAGAAAGTCAACAGATATATCC 740
Qy 81 TyTrtleuSerArgAsnGlnThrValValGlnuTyThrsAlaPheTyshArpThrArp 100
Dh 741 TACACTTTCAGAGGATCAGACTGTGTGTGTGAGATACACATATAGATAGGATAGGAT 800
Qy 101 MetPheGlnuValGlyArgSerThrGlnuSerProIleAspPheValTyThrsArpThrIle 120
Dh 801 ATGTTTCAGGTGGCGAGATCAACAGAAAGCCATATGACTTGTGTCAACAGACAT 860
Qy 121 SerGlySerGlnAsnThrArgGlnuAlaGlnIleThrsGlnSerThrIleSerArgPheAla 140
Dh 861 TCGGACAGCCAGAACCGAGCGAAAGCCAGATCAGACAGAGCAACATTCAGATTCCGC 920
Qy 141 CyseArgIleValCyseArgAsnGlnProTyThrsAlaArgIlePheAlaIshTyPhe 160
Dh 921 TGCAGATCTGTGTGCACAGAGATGAACCTTACACAGCAGGATATTCGCGCGGATTT 980
Qy 161 ArgSerSerlyshAsnIlePheleuGlyGlnuValAlaIshTyThrsAsnProArgGly 180
Dh 981 GACTCTTCCAAAACATATTTCTTGAGAAAGGCGCAAGATGTGAAGAAACCCGACGCG 1040
Qy 181 HisMetAspGlyleuThrThrsAsnGlyValleuValMetshArpArgGlyTyPheThr 200
Dh 1041 CACATGATGGGTCACTACTATGTGGTGTGTGTATGATGATCAGAGGGGCTTCACC 1100
Qy 201 GlnGlnuSerGlnProGlyValITPArgGlnuIleSerValCyseGlyshArpValTyThrsLeu 220
Dh 1101 GAGAGATCCAGACCGCGGTCTGCGCGAGATCTGTGTGTGAGATGTGTACACTTG 1160
Qy 221 ArgGlnThrArgSerlyshGlnuArgGlyshLeuValGlnuSerGlnuThrsAsnValleu 240
Dh 1161 CGAAGAACAGGTGGGCCGACAGAACAGAAAGCTGTGAAAGTGAACCACTCTCTG 1220
Qy 241 GlnAspGlySerleuIleAspLeuCyseGlyAlaThrsleuLeuTrpArgThrsAlaAspGly 260
Dh 1221 CAGAGCGGCTCCCTCATTTGACCTGTGTGGGGCCACTCTCTGTGAGAACAGCATGGG 1280
Qy 261 LeuphenisthrProthrnglnlyshlelleglnaleuargnglnuileanala 280
Dh 1281 CTTTTCATCTCCACTCGAACACATGAAACCTTCGGGAGGAGATTTAACGGCCG 1340
Qy 281 ArgProGlnCyseProValGlyLeuasnThrleuAlaPheProserlleaamrglysglu 300
Dh 1341 CGGCGCTCAGTGTCTGTGGGGCTCAACACCTGGCTTCCCGACATCAACAGAAAGAG 1400
Qy 301 ValValGlnGlnlysglnProtrPalaTytleuSerCyseGlyshValshlellyThrs 320
Dh 1401 GTGTGTGAGAGAGAGAGCCCTGGGCAATCTCACTGTGTGGCCACGTCGACGGGTAC 1460
Qy 321 AsnTrpGlyshArgSerArpThrGlnAlaAsnGlnuArgGluCyseProMetCyseArgThr 340
Dh 1461 AACGGGGCCATCGAGATGACACGAGGCCAACAGAGGAGTGTCCCATGTGCAGACT 1520
Qy 341 ValGlyProTyValProleuTrpLeuGlyCyseGlnuAlaGlyPheTyValAspAlaGly 360
Dh 1521 GTGGGCGCCATATGTGCTCTGTGGCTTGGCTGTGAGCAGGATTTTATGACGACGGA 1580
Qy 361 ProProThrshAlaPheThrProCyseGlyshValCyseSerGlnlySerAlaIshTy 380

```

Db      1581  CGCCCAACCATGCTTCACTCCCTGCGACAGTGTGCTCGGAGAAAGTCGCAAAATAC 1640
Qy      381  TTPSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
Db      1641  TGGTCTCAGATCCCGCTGCTCATGGAACCAATGCAATTCACCTGCTGCTTCTGT 1700
Qy      401  AlaThrGlnLeuValGlyGlnGlnAanCysIleLeuLeuIlePheGlnGlyProIleAap 420
Db      1701  GCTACACAGCTGAGTGGGAGCAAAACTGATCAATAATTATTTTCCAAAGSTCCAAATTGAC 1760

RESULT 5
US-09-843-905A-7
; Sequence 7, Application US/09843905A
; Patent No. US20020168683A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-843-905A-7

Alignment Scores:
Pred. No.: 5.32e-254 Length: 1263
Score: 2283.00 Matches: 419
Percent Similarity: 99.8% Conservative: 1
Best Local Similarity: 99.8% Mismatches: 0
Query Match: 99.7% Indels: 0
DB: 3 Gaps: 0

US-10-041-030-4 (1-420) x US-09-843-905A-7 (1-1263)
Qy      1  MetPheSerProGlnGlnGlnGlnIleCysAlaProAsnIleGlnProValIleTyrGly 20
Db      1  ATGTCTTCCCTGCGGAGGAGAACTGCGCCCAATAGAGAGCGGAAATACGGG 60
Qy      21  GlnLeuValIleGlnGlyTyrAanGlyAlaLeuProAsnIleAapArgGlyAArgGly 40
Db      61  GAGCTGTGTGCTCGGAGTCAATGTGTCTTACCCAAATGAGATGAGAGCGAGGAAA 120
Qy      41  SerArgPheAlaLeuTyrIleAArgProIleAAsnGlyValIleYsProSerThrValHis 60
Db      121  AGTAGATTTCCTCTACAGCGGCGCCCAAGCAAAATGTGTCAAAACCAACACGCTCAT 180
Qy      61  ValIleSerThrProGlnAlaSerIleValAlaIleSerCysLeuGlyGlnHisSerIleSer 80
Db      181  GTGATATCCACGCGCCCGGAGCATCCAGGCTATCAGCTGCAAAAGTCAACACAGTATATCC 240
Qy      81  TyrThrLeuSerArgAanGlnThrValIleValIleGlyTyrThrHisAapIleAapThrAap 100
Db      241  TACACTTTGTCCAGAGATTCAGACTGTGTGTGTGAGTACACATGATAGATACGAT 300
Qy      101  MetPheGlnValGlyAArgSerThrGlnSerProIleAapPheValIleThrAapThrIle 120
Db      301  ATGTTTCAGGTGGGAGATCAACAGAAAGCCCTATGACTTCGTGTGCACAGACAGATT 360
Qy      121  SerGlySerGlnAanThrAapGlnIleGlnIleThrGlnSerThrIleSerArgPheAla 140
Db      361  TCTGGCAGCCAGAACAGAGAGCCAGATCAACAGAGACCAATATCCAGGTTCCGCC 420
Qy      141  CysArgIleValIleCysAapArgAanGlnProTyrThrAlaArgIlePheAlaIleGlyPhe 160
Db      421  TGCAGATCTGTGTGCAAGAGAAATGAACCTTACACAGACGAGATATTCGCGCGGATTT 480

```

```

Qy      161  AapSerSerIleAAsnIlePheLeuGlyGlnValAlaAlaIleTyrIleAAsnProAapGly 180
Db      481  GACTCTTCCAAAACATATTTCTTGAGATGAAGGACGCAATGTGAAAAACCCCGACGGC 540
Qy      181  HisMetAapGlyLeuThrThrAanGlyValLeuValMetHisProAapGlyGlyPheThr 200
Db      541  CACATGAGATGGGCTCATCTAATGGCGTCTGGATGATCATCCAGAGGGGCTTCAAC 600
Qy      201  GlnGlnSerGlnProGlyValIlePArgGlnIleSerValCysGlyAapValIleThrLeu 220
Db      601  GAGAGTCCACAGCCGGGCTGTGGCGGAGATCTGTGTGTGAGATGTGACACCTTG 660
Qy      221  ArgGlnThrArgSerAlaGlnGlnAanArgGlyIleValIleGlnSerGlnThrAanValLeu 240
Db      661  CGAGAAACAGATCGGCGCCAGAACAGAGAAAGCTGTGTGAAGATGACACACGCTTG 720
Qy      241  GlnAapGlySerLeuIleAapLeuCysGlyAlaThrLeuLeuTyrAArgThrAlaAapGly 260
Db      721  CAGAGCGCTCCCTCATTTGACCTGTGTGGGCGCACTCTCTGTGAGAACACAGATGGG 780
Qy      261  LeuPheHisThrProThrGlnIleHisIleGlnAlaLeuArgGlnGlnIleAanAlaAla 280
Db      781  CTTTTCATATCTCAACTCAGAGACACATAGAAAGCTTCGCGCAGAGATTAACCGCGCC 840
Qy      281  ArgProGlnCysProValGlyLeuAanThrIleValAlaPheProSerIleAanArgIleGln 300
Db      841  CGGCTCAGATGTCTGTGTGGGCTCAACACCTGCGCTTCCAGACATCAACAGAAAGAG 900
Qy      301  ValIleGlnGlyLeuGlnProTyrAlaTyrLeuSerCysGlyHisValHisIleGlyThrHis 320
Db      901  GTGTGTGAGAGAAAGAGAGCCCTGCGCATATCTCAGTGTGGCCACGTCGACGGATCCAC 960
Qy      321  AanTyrGlyHisAArgSerAapThrGlnAlaAanGlnAArgGlnCysProMetCysAArgThr 340
Db      961  AACTGGGAGCATCGAGGTACACGAGGCGCAACAGAGAGGAGATGTCCCATGTGACAGACT 1020
Qy      341  ValGlyProTyrValProLeuThrLeuGlyCysGlnAlaGlyPheTyrValAapAlaGly 360
Db      1021  GTGGGCCCTATGTGCTCTCTGTGGCTGTGAGCGAGATTTTATGTAAACGACAGGA 1080
Qy      361  ProProThrHisAlaPheThrProCysGlyHisValCysSerGlnIleSerAlaIleTyr 380
Db      1081  CGCCCACTACATGCTTCACTCCCTGTGACACGTGTGTCGGAAGATGTGCAAAATAC 1140
Qy      381  TTPSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
Db      1141  TGGTCTCAGATCCCGTGTGCTCATGGAATCATGATTCACGCTGTGGCTTCTGTGT 1200
Qy      401  AlaThrGlnLeuValGlyGlnGlnAanCysIleLeuLeuIlePheGlnGlyProIleAap 420
Db      1201  GCTACACAGCTGAGTGGGAGCAAAACTGATCAATAATTATTTTCCAAAGSTCCAAATTGAC 1260

RESULT 6
US-10-317-250-7
; Sequence 7, Application US/10317250
; Publication No. US20030165945A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-B
; CURRENT APPLICATION NUMBER: US/10/317,250
; CURRENT FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-317-250-7

```

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 5,32e-254 | Length: | 1263 |
| Score: | 2283.00 | Matches: | 419 |
| Percent Similarity: | 99.8% | Conservative: | 0 |
| Best Local Similarity: | 99.8% | Mismatches: | 1 |
| Query Match: | 99.7% | Indels: | 0 |
| DB: | 6 | Gaps: | 0 |

US-10-041-030-4 (1-420) x US-10-317-250-7 (1-1263)

```
Qy 1 MetPheSerProGlyGlnGlnGlnHISCyEaLaProAsnlyPGLUPProVallyrTgLy 20
Db 1 ATGTTTCCCTGGCCAGAGAAACACTGCGCCCAATAGAGAGCCAGTGAATATACGGG 60
Qy 21 GluLeuValValleuGlyTyraEnglyAlaLeuProAsnlyAaPaRgLyArGaRgLy 40
Db 61 GAGCTGTGTGCTCGGGTACATGTGCTTACCCATGAGATGAGATAGAGAGAGAA 120
Qy 41 SerArgPheAlaLeuTyTyraRgProLyAaAsnlyAllybProSerThrValHis 60
Db 121 AGTAGATTGGCCCTCAAGCGGCCCAAGSCAAATGTGTCAAAACCAAGACCGTCAT 180
Qy 61 ValIleSerThrProGlnAlaSerLyAaIleSerCybLyGlnHisSerIleSer 80
Db 181 GTGATATCCAGCCCAAGCATCCAGGCTATCACTGCAAAAGGTCAACACAGTATATCC 240
Qy 81 TyrThrIleSerArgAaEngInThrValIValGluTyrThrHisAaPlyAaPThrAsp 100
Db 241 TACACTTTGTCAGAAATCAACACTGTGTGTGGAGTACACATGATAGAGATACGAT 300
Qy 101 MetPheGlnValGlyArGSerThrGlySerProIleAsPheValIThrAspThrIle 120
Db 301 ATGTTTCAAGGTGGGCAATCAAGAAAGCCCTATCCAGCTTGGTTCACAGACGATT 360
Qy 121 SerGlySerGlnAaThrAsPgluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 361 TCTGGCAGCCAGAAACAGAGAGAGCCAGATCAACAGACCAATATCCAGGTTGCC 420
Qy 141 CybArgIleValCybAaPaRgAaEngInuProTyThrHisAaGlyPheAlaIleGlyPhe 160
Db 421 TGCAGATTCGTGTGCAAGAGAAATGAACTTACACAGCAAGATATATGCGCGCAATT 480
Qy 161 AspSerSerLyAaAsnIlePheLeuGlyGluLyAaIleAlaIleTyRlyAaAsnProAspGly 180
Db 481 GACTCTTCCAAAAACATATTTCTTGGAGTAAAGCAGAAAGTGAAGAAACCCGACGGC 540
Qy 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProAaGlyGlyPheThr 200
Db 541 CACATGTAGTGGGCTCACTACTAATGCGCTCTGTGTGATGATCCACAGAGGGGCTTACC 600
Qy 201 GluGluSerGlnProGlyValITrPArGluIleSerValCySGlyAaPValITyrThrLeu 220
Db 601 GAGAGATCCAGCCCGGGGTGTGGCGGAGATCTCTGTGTGAGAGTGTATACCTTG 660
Qy 221 ArgGluThrArgSerAlaGlnGlnAaRgLyLeuLeuValGlySerGluThrAsnValLeu 240
Db 661 CGAABAACAGGTGGGCGCCAGCAAGAGAAAGCTGTGGAGAAATGAGAACCAAGTCTCG 720
Qy 241 GlnAaPglYserLeuIleAsPleuCySGlyAlaThrLeuLeuTyRPArGThrIleAaPglY 260
Db 721 CAGAGCGGCTCCCTCAATTGACTGTGTGGGCGCACTCTCTGAGAAACAGCAGATGGG 780
Qy 261 LeuPheHisThrProThrGlnIleHisIleGlnAlaLeuAaRgGlnIleAsnAlaIle 280
Db 781 CTTTTCATCTCAACTCAAGAGACATAGAAAGCCCTCGGCGAGAGATTAAAGCGCGCC 840
Qy 281 ArgProGlnCybProValGlyLeuAaAsnThrLeuAlaPheProSerIleAsnArgLyGlu 300
Db 841 CGGCGTCAAGTCTCTGTGGGCTCAACACCTGCGCTTCCCAAGCATCAAGAGAAAGAG 900
Qy 301 ValIValGluGluTyRgInProITrPAIATyrLeuSerCySGlyHisIleGlyTyRHis 320
Db 901 GTGGTGGAGAGAGAGCCCTGGGCAATATCTCAATTGTGGCCACGTGACAGGATACAC 960
```

```
Qy 321 AsnITrGlyHisArgSerAaPThrGlnAlaAsnGluAaRgGluCybProMetCybArgThr 340
Db 961 AACTGGGGCCATGTGAGTGAACGAGAGCCAAAGAGAGAGTGTCCATGTGCAAGACT 1020
Qy 341 ValGlyProTyRValProLeuITrPLeuGlyCySGluAlaGlyPheTyRValAaPAlaGly 360
Db 1021 GTGGGCGCCCTATGTGCTCTCTGTGGCTGTGAGAGCAGATTTATGTAGACGAGGA 1080
Qy 361 ProProThrHisAlaPheThrProCySGlyHisIleValCySGerGlyLybSerAlaLyTyR 380
Db 1081 CCGCAACTATATCTTCACTCTCTGTGACAGCTGTGCTCGAGAAATGTGCAAAATAC 1140
Qy 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaIaCybProPheCyS 400
Db 1141 TGGTCTCAATCCCGTTGGCTCATGAACTCAATGCAATTCAGCTGCTGCTTCTGT 1200
Qy 401 AlaITrGlnLeuValGlyGlnAaAsnCybIleLybLeuIlePheGlnGlyProIleAaP 420
Db 1201 GCTACACAGCTGTGGGAGCAAAACTGCATCAATTAATTTCCAAAGGTCAATTGAC 1260
```

RESULT 7

US-10-258-703-7
; Sequence 7, Application US/10258703
; Publication No. US20040034199A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; TITLE OR INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-US
; CURRENT APPLICATION NUMBER: US/10/258,703
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 7
; LENGTH: 1263
; TYPE: DNA
; ORGANISM: Homo sapiens
US-10-258-703-7

| | | | |
|------------------------|-----------|---------------|------|
| Alignment Scores: | | | |
| Pred. No.: | 5,32e-254 | Length: | 1263 |
| Score: | 2283.00 | Matches: | 419 |
| Percent Similarity: | 99.8% | Conservative: | 0 |
| Best Local Similarity: | 99.8% | Mismatches: | 1 |
| Query Match: | 99.7% | Indels: | 0 |
| DB: | 7 | Gaps: | 0 |

US-10-041-030-4 (1-420) x US-10-258-703-7 (1-1263)

```
Qy 1 MetPheSerProGlyGlnGlnGlnHISCyEaLaProAsnlyPGLUPProVallyrTgLy 20
Db 1 ATGTTTCCCTGGCCAGAGAAACACTGCGCCCAATAGAGAGCCAGTGAATATACGGG 60
Qy 21 GluLeuValValleuGlyTyraEnglyAlaLeuProAsnlyAaPaRgLyArGaRgLy 40
Db 61 GAGCTGTGTGCTCGGGTACATGTGCTTACCCATGAGATGAGATAGAGAGAGAA 120
Qy 41 SerArgPheAlaLeuTyTyraRgProLyAaAsnlyAllybProSerThrValHis 60
Db 121 AGTAGATTGGCCCTCAAGCGGCCCAAGSCAAATGTGTCAAAACCAAGACCGTCAT 180
Qy 61 ValIleSerThrProGlnAlaSerLyAaIleSerCybLyGlnHisSerIleSer 80
Db 181 GTGATATCCAGCCCAAGCATCCAGGCTATCACTGCAAAAGGTCAACACAGTATATCC 240
Qy 81 TyrThrIleSerArgAaEngInThrValIValGluTyrThrHisAaPlyAaPThrAsp 100
Db 241 TACACTTTGTCAGAAATCAACACTGTGTGTGGAGTACACATGATAGAGATACGAT 300
Qy 101 MetPheGlnValGlyArGSerThrGlySerProIleAsPheValIThrAspThrIle 120
Db 301 ATGTTTCAAGGTGGGCAATCAAGAAAGCCCTATCCAGCTTGGTTCACAGACGATT 360
```


| | | | |
|--|------|--|------|
| QY | 121 | SeelIySerGIaAntThrAspGIaValAGInIleThrGInSerTrIleSeArGIpHea | 140 |
| Db | 361 | TCGGGACCCAGAACACGGAGCAAGACCAGATCACAGAGACCACTATCCAGTTGCC | 420 |
| QY | 141 | CysArgIIeValCyAspArgAsnGluProTyrrThralaArgIIleHeaIlaIGIyPe | 160 |
| Db | 421 | TGCAGGATCGGTGCGACAGGAATGAACCTTACACAGCAGGATATTGCGCGGANTTY | 480 |
| QY | 161 | AspSerSerIySaenIIlePheLeuGIyGluIySaIaIySTpIySaenProAspGIy | 180 |
| Db | 481 | GACTCTTCCAAAACATATTTCTTGAGATAAGGAGCAAAAGTGAATAAATCCCGACGGC | 540 |
| QY | 181 | HisMetAspGIyLeuThrThrAsnGIyValIeuValMetHisProArgGIyGIyPheThr | 200 |
| Db | 541 | CACATGATGGGCTCATCACTAAATGGCCTCTGGTGAATGATCCACAGGGGGCTTCAAC | 600 |
| QY | 201 | GIuGIuSerGIaProGIyValITrpArgGIuIleSerValCySGIyAspValTyrrThreIu | 220 |
| Db | 601 | GAGAGATCCAGCCCGGGGCTGGCGCAGAGATCTCTGCTGGAGATGTGTACACCTTG | 660 |
| QY | 221 | ArgGIuThrArgSerAlaGIaGIaArgGIyIyLeuValGIuSerGIuThrAspValIeu | 240 |
| Db | 661 | CGAGAAACACAGGTGGCCCGACAGCAACAGAAAGCTGTGAAGTGAACACCAAGTCTTG | 720 |
| QY | 241 | GIaAspGIySerIleIIeAspLeuCySGIyValaThrIleuLeuTrpArgThralaAspGIy | 260 |
| Db | 721 | CAGAGCGGCTCCCTCATTTAGCACTGTGTGGGCCACTCTCTCTGGAGAACAGCAGATGG | 780 |
| QY | 261 | LeuPheHisThrProThrGIuIyHisIIleGIuIaIeuArgGIuIIleAsnAlaIa | 280 |
| Db | 781 | CTTTTTCATATCCCAACTCAGAACACATAGAACCCCTCGCGAGGATTTAAGCGGCC | 840 |
| QY | 281 | ArgProGIuCySProValGIyIleuAsnThreIuIaPheProSerIIeAsnArgGIu | 300 |
| Db | 841 | CGGGCTCAGTGTCTCGTGGGGCTCAACCCCTGGGCTTCCCGAGATCAACAGAAAGAG | 900 |
| QY | 301 | ValValGIuGIuIySGInProTrpAlaTyrrLeuSerCySGIyHisValIIleGIyThiS | 320 |
| Db | 901 | GTGGTGGAGAGGAAGACAGCCCTGGGGCANATCTCACTGTGGCCACGTGCACGGTTACAC | 960 |
| QY | 321 | AsnTrpGIyHisArgSerAspThrGIuIaIaenGIuArgGIuCySPrometCyArgThr | 340 |
| Db | 961 | AACGGGGCCATCGGAGTGACACGGAGGCCAAACAGAGGGAGTGTCCCATGTGACAGACT | 1020 |
| QY | 341 | ValGIyProTyrrValProLeuTrpLeuGIyCySGIuIaGIyPheTyrrValAspAlaGIy | 360 |
| Db | 1021 | GTGGGCCCCCTTAATGGCTCTCTGGCTTGGCTGTAGGAGGAAATTTTATGTAAGCAGAGA | 1080 |
| QY | 361 | ProProThrHisAlaPheThrProCySGIyHisValCySerGIuIySerAlaIyTyrr | 380 |
| Db | 1081 | CCGCAACTCATGCTTTCACCTCCGTGTGACACCGTGTCTCGAGAACTCGCAAAATAC | 1140 |
| QY | 381 | TrpSerGIuIIleProLeuProHisGIyThrHisAlaPheHisAlaIaCySPrometCyS | 400 |
| Db | 1141 | TGGTCTCAGATCCCGTGGCTCTCATAGGAATCATGATTCACGCTGCTGCCCTTCTGCT | 1200 |
| QY | 401 | AlaThrGIuIeuValaGIyGIuGIaAsnCyTrIIeIySerIlePheGIuIyProIIeAsp | 420 |
| Db | 1201 | GCTACACAGCTGTGTGGGAGGAAACTGATCAATTAATTTTCAAGGTCAATTTGAC | 1260 |
| RESULT 8 | | | |
| US-09-843-905A-5 | | | |
| Sequence 5, Application US/09843905A | | | |
| Patent No. US20020168683A1 | | | |
| GENERAL INFORMATION: | | | |
| APPLICANT: Bird, Timothy A. | | | |
| APPLICANT: Cosman, David J. | | | |
| TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES | | | |
| FILE REFERENCE: 2990-A | | | |
| CURRENT APPLICATION NUMBER: US/09/843,905A | | | |
| CURRENT FILING DATE: 2001-04-27 | | | |
| PRIOR APPLICATION NUMBER: US 60/200,198 | | | |

```

: PRIOR FILING DATE: 2000-04-28
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 5
: LENGTH: 1260
: TYPE: DNA
: ORGANISM: Mus musculus
US-03-843-905A-5

Alignment Scores:
Pred. No.: 1,566-241 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.48 Conservative: 10
Best Local Similarity: 95.08 Mismatches: 10
Query Match: 95.08 Indels: 1
DB: 3 Gaps: 1

US-10-041-030-4 (1-420) * US-09-843-905A-5 (1-1260)

QY 1 MetPheSerProGlyGlnGluHisCybaLaProAenLySerGlnProValLySerGly 20
Db 1 ATGTTTTCCTCCGGGCGCAGAGGAAACCCAGCGCCCTCCCAACAGACCCGGTGAATACGGG 60

QY 21 GluLeuValValLeuGlyTyraSmGlyValaLeuProAenGlyAspArgGlyArgArgLys 40
Db 61 GAGCTGGTGTCTCTGGGGTCAATGTGTCTTACTTAATGTGTACAGGGGCGAGAGGAA 120

QY 41 SerArgPhealaLeuTyryaGrProLySaLaenGlyValaLySerPheSerThrValHis 60
Db 121 AGCAGATTTCCTCTTAATAGGGACCTACGCCACTGTGTCAAAACCCAGCACAATCCAC 180

QY 61 ValLeuSerThrProGlnAlaSerLySaLaLeuSerCybLyGlyGlnHisSerLLeuSer 80
Db 181 ATGGCTTCCACACCAACAGCGCTCCAGGCCATCACTCCAGAGGACATCAAGACATATG 240

QY 81 TyrThrLeuSerArgAsnGlnThrValaValaGlyTyrrThrHisAspLySaSerThrAsp 100
Db 241 TACACGTGTGTCACGGAGCGACAGCGTAGTGGAGTACACACGATTAAGACACGAC 300

QY 101 MetPheGlnValaGlyArgSerThrGluSerProLLeaspPheValaValThrAspThrLe 120
Db 301 ATGTTTCAAGTGGGAGGATCAACAGAAAGCCCATTAAGCTTGTGTGTCAACAGACGGTT 360

QY 121 SerGlySerGlnaSerThrAspGlnaLacGlnLeuThrGlnSerThrLLeSerArgPheala 140
Db 361 TCCGGCGGTCCAGAACGAAGAT---GCCCAAGATCACACAGACGACCATCTCTAGTTCGGA 417

QY 141 CysArgGlnLeuValCybaAspArgAsnGluProTyrrThrAlaArgGlnPhealaLacGlyPhe 160
Db 418 TGCAGGATCGTGTGTGACAGAAACAGGCATATACAGACGCAATATTCGGCGAGGATTC 477

QY 161 AspSerSerLySaenLLePheLeuGlyGlnLySaLaLaLysTyrrLySaenProAspGly 180
Db 478 GATCTTCCAAAANAATCTTCTTCTTGAGAGAAAGACGCAAAATGAGAAAACCTGATGGA 537

QY 181 HisNecAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db 538 CACATGATGAGCATCATCAACCAATGTGTCTTAAGTAGACACCCCAAGAGAGGCTTCACC 597

QY 201 GluGluSerGlnProGlyValaITPArgGlnLLeSerValCybGlyAspAlaTyrrThrLeu 220
Db 598 GAGGATATCCAGCCTGAGATCTGAGAGAGAGANTCTGTCTGTGGGAAATGTACACTTG 657

QY 221 ArgGlnThrArgSerAlaGlnGlnArgGlyLyLeuValaGluSerGlyThrAsnValLeu 240
Db 658 CGAGAGACCAAGTCCGCCACAGACAGGGGAAAGCTGTGTGAAGATGAGACCAAGCTCTG 717

QY 241 GlnAspGlySerLeuLLeAspLeuCybGlyValaThrLeuLeuTyrrArgThrLLeAspGly 260
Db 718 CAAGACGGCTCCCTATGACCTGTGTGGGGCACTCTCTGTGAGAAACCCAGATGGC 777

QY 261 LeuPheHisThrProThrGlnLyHisLLeGlnValaLeuArgGlnGlnLLeAsnAlaLa 280

```

Db 778 CTTTTCACGCTCTCTACTCAGAAACATGAGACCTCCGGCAGAGATCAATGACAGCC 837
Qy 281 ATGProGlnCysProValGlyLeuAanThrLeuAlaPheProSerIleAanArgValGlu 300
Db 838 CGACCCAGAGGCCCCGGGGCTTAAACACCTGGCCTTCCCAATCAACCGAAGGAA 897
Qy 301 ValValGluGlyValGlnProTPrAlaTyrLeuSerCysGlyValAlaHisGlyTyrHis 320
Db 898 GTGGTGGAAGAAGACGAGCCCTGGGCAATCCTGAGCTGGCCGCAATGACAGCTACAC 957
Qy 321 AsnTPrGlyHisArgSerAapThrGluAlaAanGluArgLysCysProMetCysArgThr 340
Db 958 AGCTGGGGCCATCGAGCGAAGCGGAAAGCCAAACGAGGAGGAGTCTCCATGTGAGGACT 1017
Qy 341 ValGlyProTyrValProLeuTPrLeuGlyCysGluAlaGlyPheTyrValAlaPheAlaGly 360
Db 1018 GTGGGCCCCCTACGCTCTCTGAGCTGGAGGAGGAGATTTTAATGCGAAGCGGGA 1077
Qy 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr 380
Db 1078 CCCCCAACTCAGCCTTTCACCCCCCTGGGGCAGCTGTTCAGAAAGTCTGCCAAGTAC 1137
Qy 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
Db 1138 TGGTCGAGATCCCACTGCGCCCAAGCAAGCAGCGCTTTCATGCGGCTGTCCGTTCTGC 1197
Qy 401 AlaThrGlnLeuValGlyGluAanCysIleLysLeuIlePheGlnGlyProIleAap 420
Db 1198 GCCACGAGCTGGTGGTGAACAGAACTGCAATTAATTTTCCAAAGTCCAGTGAGC 1257

RESULT 9
US-10-317-250-5
; Sequence 5, Application US/10317250
; Publication No. US20030165945A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; APPLICANT: Li, Xiaoxia
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2890-B
; CURRENT APPLICATION NUMBER: US/10/317,250
; CURRENT FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-317-250-5

Alignment Scores:
Pred. No.: 1,566-241 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.4% Conservative: 10
Best Local Similarity: 95.0% Mismatches: 10
Query Match: 95.0% Indels: 1
Gaps: 1

US-10-041-030-4 (1-420) x US-10-317-250-5 (1-1260)

Qy 1 MetPheSerProGlyGlnGluGluHisCysAlaPheAanLysGluProValLysTyrGly 20
Db 1 ATGTTTCCCCGGGCGAGAGAAACCAAGCCCCCAACAGAGCCGGTGAATTAAGGG 60
Qy 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAanGlyAapArgLysArgArgLys 40
Db 61 GAGCTGTGCTCGGGGTACAATGTGCTTACCTAATGTGTGACAGGGGAGAGGAAA 120
Qy 41 SerArgPheAlaLeuTyrLysArgProLysAlaAanGlyValLysProSerThrValHis 60
Db 121 AGCAGATTGTGCTCTTAAGCGGAGCTTAACGCTGTGTGCAAAACCAAGCAATATCAC 180
Qy 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80

Db 181 ATGTCCTCCACACCAAGGGGCTCAAGGCCATGAGCTCCAGAGACATCAACACATATCG 240
Qy 81 TyrThrLeuSerArgAanGlnThrValValGluTyrThrHisAanPheAapThrAap 100
Db 241 TACACGTTGTCAAGGAGCCAGACGATAGTGTGAATACACACATTAAGACAGGAC 300
Qy 101 MetPheGlnValGlyArgSerThrGlySerProIleAapPheValValThrAapThrIle 120
Db 301 ATGTTTCAGGTGGGAGGTCAACAGAAAGCCCAATTAATGCTGTGCTGACAGACGGTT 360
Qy 121 SerGlySerGlnAanThrAapGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 361 TCCGGCGGTCAAGACGAAGAT--GCCAGATACACAGAGACCAATCTCTAGTTTCGCA 417
Qy 141 CysArgLysValCysAapArgAanGluProTyrThrAlaArgLysPheAlaIleGlyPhe 160
Db 418 TGCAGATCTGTGTACAGAAAGCAGCCATATACAGCAGCATATTCGGGAGGATTC 477
Qy 161 AspSerSerLysAanIlePheLeuGlyGluLysAlaAlaLysTPrLysAanProAapGly 180
Db 478 GATTCCTCCAAAATATCTTCTTGGAGAGAAAGCAGAAATGGAATAAAGCTTGATGA 537
Qy 181 HisMetAapGlyLeuThrThrAanGlyValLeuValMetHisAProArgGlyGlyPheThr 200
Db 538 CACATGATGATGACTCACTACCAATGTGTCTGATGATGACCCGCAAGAGGACTTCACC 597
Qy 201 GlnGluSerGlnProGlyValTPrArgGluIleSerValCysGlyLysAapValTyrThrLeu 220
Db 598 GAGGAATCCAGGCTGTGAGCTGAGAGAGATCTGTGTGTGGAGATGTGACACCTTG 657
Qy 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGlyLysSerGluThrAanValLeu 240
Db 658 CGAGACACAGCTCTGGCCAGCAGAGGAGGAAAGCTGTGAAGTGAAGCAACGCTCTG 717
Qy 241 GlnAapGlySerLeuIleAapLeuCysGlyAlaThrLeuLeuTPrArgThrAlaAapGly 260
Db 718 CAAGACGGCTCCTCAATTAATGCTGTGGGGCCACTCTCTCGAAGAACCGCAGATGGC 777
Qy 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuAanGlnGluIleAanAla 280
Db 778 CTTTTCACGCTCTCTCTCAAGAAACATTAAGAACCTCCGGGAGAGATCAATGACGCC 837
Qy 281 ArgProGlnCysProValGlyLeuAanThrLeuAlaPheProSerIleAanArgLysGlu 300
Db 838 CGACCCAGTGCCTTTCACCTTAAACCTTGGCTTCCCAAGCATCAACCGAAGAA 897
Qy 301 ValValGluGlyValGlnProTPrAlaTyrLeuSerCysGlyValAlaHisGlyTyrHis 320
Db 898 GTGGTGGAAGAAGACGAGCCCTGGGCAATCCTGAGCTGGCCGCAATGACAGCTACAC 957
Qy 321 AsnTPrGlyHisArgSerAapThrGluAlaAanGluArgLysCysProMetCysArgThr 340
Db 958 AGCTGGGGCCATCGAGCGAAGCGGAAAGCCAAACGAGGAGGAGTCTCCATGTGAGGACT 1017
Qy 341 ValGlyProTyrValProLeuTPrLeuGlyCysGluAlaGlyPheTyrValAlaPheAlaGly 360
Db 1018 GTGGGCCCCCTACGCTCTCTGAGCTGGAGGAGGAGATTTTAATGCGAAGCGGGA 1077
Qy 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr 380
Db 1078 CCCCCAACTCAGCCTTTCACCCCCCTGGGGCAGCTGTTCAGAAAGTCTGCCAAGTAC 1137
Qy 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
Db 1138 TGGTCGAGATCCCACTGCGCCCAAGCAAGCAGCGCTTTCATGCGGCTGTCCGTTCTGC 1197
Qy 401 AlaThrGlnLeuValGlyGluAanCysIleLysLeuIlePheGlnGlyProIleAap 420
Db 1198 GCCACGAGCTGGTGGTGAACAGAACTGCAATTAATTTTCCAAAGTCCAGTGAGC 1257

RESULT 10
US-10-085-117-285

```

; Sequence 285, Application US/10085117
; Publication No. US2003023234A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER
; FILE REFERENCE: 529452000121
; CURRENT APPLICATION NUMBER: US/10/085,117
; CURRENT FILING DATE: 2002-02-27
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 361
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 285
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-085-117-285

```

```

Alignment Scores:
Pred. No.: 1.56e-241 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.4% Conservative: 10
Best Local Similarity: 95.0% Mismatches: 10
Query Match: 95.0% Indels: 1
DB: 6 Gaps: 1
US-10-041-030-4 (1-420) x US-10-085-117-285 (1-1260)

```

```

QY 1 MetPheSerProGlyGlnGluGlnHisCysAlaProAsnLysGluProValLysTYRGLY 20
DB 1 ATGTTTCCCCGGGCGGAGGAGAACCCAGCGCCCCCAACAGAGCCGGTGAATACGGG 60
QY 21 GluLeuValLeuGlyTYRAsnGlyAlaLeuProAsnLysAspArgLysArgLys 40
DB 61 GACTGGGTGCTCTGGGGGTCAATAGTGCTTACCTTAATGATGTCACAGGGCAGAGAGAA 120
QY 41 SerArgPheAlaLeuTYRValArgProLysAlaAsnGlyValLysProSerThrValHis 60
DB 121 AGCAGATTGCCCCCTTAATAGCGGACCTACGCGAGTGCTCAAAACCCAGACAAATCCAC 180
QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80
DB 181 ATGCTCTCCACACACAGCGGCTCCAGGCACTACGCTCCAGGACATCCACAGATATCG 240
QY 81 TYRThrLeuSerArgAsnGlnThrValAlaGluTYRHisAspLysAspThrAsp 100
DB 241 TACACGTTGTCACAGGACGACGAGTGCTGTGAGTACACACGATTAAGACACGAGC 360
QY 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValValThrAspThrIle 120
DB 301 ATGTTTCAGGTGGGCGAGTCAACGAAGCCCATTCATCTCGGTCAACAGACACGGTT 360
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 361 TCCGGCGGTCAAGCAAGAT--GCCCATCAACACAGACCACTCTAGGTTGCA 417
QY 141 CysArgIleValCysAspArgAsnGluProTYRThrAlaArgIlePheAlaIleGlyPhe 160
DB 418 TGCAGATCTCTGTGACAGAAACGACCATATACAGCAGCATATTCGGCGAGATTC 477
QY 161 AspSerSerLysAsnIlePheLeuGlyGluLysAlaAlaLysTYRAsnProAspGly 180
DB 478 GATTTCTTCAAAATATATCTTTCTTGAGAGAAAGACAAATGAGAAACCTGATGGA 537
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProAspArgLysGlyPheThr 200
DB 538 CACATGATGACTCACTACCAATGCTCTAGTAGATGACCCGCAAGAGGCTTCACC 597
QY 201 GluGlnSerGlnProGlyValLysArgGluIleSerValCysGlyAspValTYRThrLeu 220
DB 598 GAGGAAATCCAGCGCTGAGTCTGAGAGAGATCTCTCTGTGGATGTGTACACCTTG 657

```

```

QY 221 ArgLutThrArgSerAlaGlnGlnArgLysLeuValGluSerGluThrAsnValLeu 240
DB 658 CGAGAGACCAAGGTGCGCCACGACGAGGGAGAAAGCTGTGGAATGAGACCAACGCTCTG 717
QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTrpArgThrAlaAspGly 260
DB 718 CAAAGCGGCTCCCTCATTAACGCTGTGGGGCAGCACTCTCTGTGAGAAACCGCAGATGGC 777
QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnIleAlaAla 280
DB 778 CTTTTCAGCTCTCTCTCAGAGCAATAGAGCCCTCCGCGAGAGATCAATGACGCC 837
QY 281 ArgProGluCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
DB 838 CAGCCCAAGTCCCGCTGGGCTTAAACCTTGCTGCTTCCAGCATCAACCGAAGAA 897
QY 301 ValValGluGlnLysGlnProTrpAlaTYRLeuSerCysGlyHisValHisGlyTYRHis 320
DB 898 GTGTGGAAGAGAGAGAGCCCTGGGCTATGAGCTGGGCGCATGTGACCGCTTCCAC 957
QY 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgLysCysProMetCysArgThr 340
DB 958 AGCTGGGGCCATCGAGCGACCGGAAAGCCAAAGAGGAGAGTGTCCATGTGCAGAGCT 1017
QY 341 ValGlyProTYRValProLeuTrpLeuGlyCysGluAlaGlyPheTYRValAspAlaGly 360
DB 1018 GTGGGCCCCCTACCTCTCTGTGGCTGTGAGGCGAGATTATATGATGCGGGA 1077
QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTYR 380
DB 1078 CCCCACACTCAGCCTTACCCCTGGGCGACGTGTGTCAGAAAGCTGCGCAAGTAC 1137
QY 381 TrpSerGlnIleProLeuProHisGlyTYRHisAlaPheHisAlaAlaCysProPheCys 400
DB 1138 TGTGTGCAATGTCACATGCCCCACGAGACGACGCGCTTCATGCGCCGTGCTGTCG 1197
QY 401 AlaThrGlnLeuValGlyGlnGlnAsnCysIleLeuLeuIlePheGlnIleProIleAsp 420
DB 1198 GCCACGACGTGTTGTGTAACGAACCTGATCAATGATTTTCCAGAGTCCAGTGAGC 1257

```

RESULT 11

```

US-10-258-703-5
; Sequence 5, Application US/10258703
; Publication No. US2004003419A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN BELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-US
; CURRENT APPLICATION NUMBER: US/10/258,703
; CURRENT FILING DATE: 2002-10-24
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 5
; LENGTH: 1260
; TYPE: DNA
; ORGANISM: Mus musculus
; US-10-258-703-5

```

Alignment Scores:

```

Pred. No.: 1.56e-241 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.4% Conservative: 10
Best Local Similarity: 95.0% Mismatches: 10
Query Match: 95.0% Indels: 1
DB: 7 Gaps: 1
US-10-041-030-4 (1-420) x US-10-258-703-5 (1-1260)

```

```

QY 1 MetPheSerProGlyGlnGluGlnHisCysAlaProAsnLysGluProValLysTYRGLY 20
DB 1 ATGTTTCCCCGGGCGGAGGAGAACCCAGCGCCCCCAACAGAGCCGGTGAATACGGG 60

```

```

Qy 21 GluLeuValIleuGlyTYraNGlyValAleuProaNGlyAaPaRGlyAaRGlyA 40
Db 61 GAGCTGGTGGTCCGTGGGGAACAATGGTCTTACTTAATGGTGAACAAGGGAAGAAA 120
Qy 41 SerArghealAleuTYlysaNGProlysaAlaNGlyValIysProSerThrValHis 60
Db 121 AGCAGATTGCCCCCTTAAGCCGACCTACGCCGCTGGTGTCAAAACCAGACCAATCCAC 180
Qy 61 ValIleSerThrProGlnAlaSerIysAlaIleSerCysIysGlyGlnHisSerIleSer 80
Db 181 ATGGTCTCCACACACAGCGGCTCCAGGCGCATCTCCAGAGGACATCCAGCATATCG 240
Qy 81 TYrThrLeuSerArgaNGInThrValIValIValIuTYrThriIsaPlysaPThrAsp 100
Db 241 TACACGTTGTCCAGGACCGACAGCGTGTGGTGAAGTACACACGATTAAGAACAGGAC 300
Qy 101 MecPheGlnValGlyAaSerThrGluSerProIleAspPheValIValIAspThrIle 120
Db 301 ATGTTTCAGGTGGGCGAGGTCAACAGAAAGCCCATTCGTCGTCCACAGACACGGTT 360
Qy 121 SerGlySerGlnaNGInThrAspGlnIleGlnIleThrGlnSerThriIleSerArgPheAla 140
Db 361 TCCGGCCGTGAAACGAAAGAT--GCCCAATGACACAGACACCATCTTGGTGGCA 417
Qy 141 CysaRGlyIleValCysaPaRGaNGluProTYrThriIaArgIlePheAlaIaGlyPhe 160
Db 418 TGCAGGATCGTGTGTGACAGGAACGAGCAATACAGACCGCATTTCCGGCGAGATTC 477
Qy 161 AspSerSerIysaNGInIlePheLeuGlyGluIysAlaIaIyTYrIyIyIyIyIyIy 180
Db 478 GATTCTTCAAAATACTCTTCTTGGAGAGAAACACCAAAATGAAAAACCCCTGATGGA 537
Qy 181 HisMetAspGlyLeuThrThraNGIyValIleuValIleThiIsaPAspRGlyGlyPheThr 200
Db 538 CACATGGATGGACTCACTCAACCAATGGTCTTGAAGTACACCGCAAGGAGGCTTCAAC 597
Qy 201 GluGluSerGlnProGlyValITrPaRGluIleSerValCysGlyAspValTYrThrLeu 220
Db 598 GAGGAATCCAGCGCTGGAGTGTGAGAGAGATCTCTGTGTGGGATGTGTACACTTG 657
Qy 221 ArgGluThraNGSerAlaGlnGlnArgGlyIyIysLeuValGluSerGluThraNGValIleu 240
Db 658 CGAAGACCGAGGTGGCGCCAGCAAGGGGAAAGCTGGGGAAGTGAACCAACGCTCGT 717
Qy 241 GluAspGlySerIleuIleAspLeuCysGlyValaThreLeuLeuTPaRGThriIaAspGly 260
Db 718 CAAGACGCTCCCTCATTTGACCTGTGGGCGCACTCTCTGAGAAACCGCAGATGGC 777
Qy 261 LeuPheHisThraNGProThriGlnIyHisIleGluIaLeuArgGlnIyIleAspAlaIa 280
Db 778 CTTTTCACGCTCTACTCAGAACCACTAAGAACCCCTCCGGCAGAGATCAATCCAGCC 837
Qy 281 ArgProGlnCysProValGlyLeuAsnThriLeuAlaPheProSerIleAsnArgIyGlu 300
Db 838 CGAACCCAGTGCCTCCGTGGGCTTAAACAACCTGGCCTTCCCAACATCAACCGAAGAA 897
Qy 301 ValIValIleGluIyysGInProThraIyIreuserysGlyHisIyValHisIyTYrHis 320
Db 898 GTGGTGAAGAGAGAGAGCCCTGGGCATACCTGAGCTGGCGCATGTGCAGGCTACAC 957
Qy 321 AsnTrpGlyHisAspSerAspThriGluIaAsnGluArgGluCysProMetCysArgThr 340
Db 958 AGCTGGGGCCATCCGAGGAGACCGGAAACCAAGAGAGAGTCTCCATGTGCAGACT 1017
Qy 341 ValGlyProTYrValProLeuTrpLeuGlyCysGluIaGlyPheTYrValIAspAlaGly 360
Db 1018 GTGGGCCCCCTAGTCCCTCTGTGGCTGGGCTGTGAGGAGAGATTTATGTGATGCGGGA 1077
Qy 361 ProProThriHisAlaPheThraNGProCysGlyHisIyValCysSerGlyIyIySerAlaIyTYr 380
Db 1078 CCCCACACTCAACGCTTTCACCCCTGGCGGACGCTGTTCAGAAATGTGCGCAAGTAC 1137
Qy 381 TrpSerGlnIleProLeuProHisGlyThriIsaIlePheHisIyAlaIaCysProPheCys 400

```

```

Db 1138 TGGTCCGATCCCACTGCCCCACGAGACGACGCGCTTTCATCGCCCTGTCTGCG 1197
Qy 401 AlaThriGluLeuValGlyGluGlnaNGCysIleIyLeuIleIlePheGlnGlyProIleAsp 420
Db 1198 GCCACGAGCTGGTGGTGAACAGAACTGCATCAATTTTCCAAAGTCCAGTGGAC 1257

RESULT 12
US-10-197-666A-133
; Sequence 133, Application US/1019766A
; Publication No. US20030092037A1
; GENERAL INFORMATION:
; APPLICANT: ASAH KASEI KABUSIKI KAISYA
; TITLE OF INVENTION: Ekl phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197,666A
; CURRENT FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 133
; LENGTH: 1717
; TYPE: DNA
; ORGANISM: Mus musculus
; NAME/KEY: Mub musculus
; FEATURE:
; LOCATION: (218)..(1474)
; NAME/KEY: unsure
; LOCATION: 1288..1301
; OTHER INFORMATION: n=A or C or G or T
US-10-197-666A-133

Alignment Scores:
Pred. No.: 6,41e-234 Length: 1717
Score: 2111.50 Matches: 388
Percent Similarity: 95.2% Conservative: 12
Best Local Similarity: 92.4% Mismatches: 19
Query Match: 92.2% Indels: 1
DB: Gaps: 1

US-10-041-030-4 (1-420) x US-10-197-666A-133 (1-1717)
Qy 1 MecPheSerProGlyGlnGluGlnHisCysAlaProAsnIyGluProValIyTYrGly 20
Db 218 ATGTTTCCCGGGCCAGAGAGAAACCGAGCGCCCAACAGAGGCGCGTGAATACAGG 277
Qy 21 GluLeuValIleuGlyTYraNGlyValAleuProaNGlyAaPaRGlyAaRGlyA 40
Db 278 GAGCTGGTGGTCCGTGGGGAACAATGGTCTTACTTAATGGTGAACAAGGGAAGAAA 337
Qy 41 SerArghealAleuTYlysaNGProlysaAlaNGlyValIysProSerThrValHis 60
Db 338 AGCAGATTGCCCCCTTAAGCCGACCTACGCCAGTGGTGTCAAAACCAGACCAATCCAC 397
Qy 61 ValIleSerThrProGlnAlaSerIysAlaIleSerCysIysGlyGlnHisSerIleSer 80
Db 398 ATGGTCTCCACACACAGCGGCTCCAGGCGCATGACCTCCAGAGGACATCAACGATATCG 457
Qy 81 TYrThrLeuSerArgaNGInThrValIValIValIuTYrThriIsaPlysaPThrAsp 100
Db 458 TACACGTTGTCCAGGACCGGAGCGGTGTGGTGAAGTACACACGATTAAGAACAGAC 517

```

```

QY 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValThrAspThrIle 120
Db 518 ATGTTTCAGGTGGGCGAGTCAACAGAAAGCCCACTTACCTTCGGTGTACAGACCGGTT 577
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 578 TCCGGGGGCGTCAAGAACGAAGAT---GCCCATGATCAACAGACGACCATCTCTAGGTTGCA 634
QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
Db 635 TGCAGAGTCTGTGTGACAGAAACGACCATACACAGCAGCATATTCGGCGACGAGATTC 694
QY 161 AspSerSerIleValIlePheLeuGlyGluValAlaIleValIlePheIleAspProAspGly 180
Db 695 GATTCTTCCAAAATATCTTTCTTGGAGAAAGACGAAATGAAATGAAATGAAATGAAATGAA 754
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db 755 CACATGATGATGACTCACTACCAACGGTGTCTGTGTATGACCCGCAAGAGGCTTACCC 814
QY 201 GlnGlnSerGlnProGlyValITPArgGluIleSerValCysGlyAspValITyrThrLeu 220
Db 815 GAGGAATCCAGCCTGAGTCTGAGAGAGATCTCTGTCTGTGGGAGTATATACACTTG 874
QY 221 ArgGluThrArgSerArgAsnGlnArgGlyValLeuValGlnSerGluThrAsnValLeu 240
Db 875 CGAGAGACCGAGTGGCGCCAGAGAGGGAAGCTGTGTGAAAGTGAAGACCAACGTTCTG 934
QY 241 GlnAspGlySerLeuIleAspLeuCysGlyValIThrLeuLeuTyrArgThrAlaAspGly 260
Db 935 CAAGACGGCTCTCTCATTTGACTGTGTGTGGGCGCATCTCTCTGTGAGAACCGCAGATGCG 994
QY 261 LeuPheHisThrProThrGlnIleValHisIleGluAlaLeuArgGlnIleAsnAla 280
Db 995 CTTTTCACGCTCTACTACAGAACACATAGAAAGCTCCGCGACGAGATCAATGACGCC 1054
QY 281 ArgProGlnCysProValIleLeuAsnThrLeuAlaPheProSerIleAsnArgValGlu 300
Db 1055 CGACCCAGAGGCCCGGTGGGCTTTAACACCTGGCTTCCCGACGATCAACCGGAAGGAA 1114
QY 301 ValValGlnGluLeuArgGlnProTyrAlaTyrLeuSerCysGlyHisValHisIleTyrHis 320
Db 1115 GTGGTGGAAAGAGACAGCCCTGGGCACTGAGCTGGCGCATGTGCACCGGCTTACAC 1174
QY 321 AsnTyrGlyHisArgSerAspThrGlnAlaAsnGluArgGluCysPrometCysArgThr 340
Db 1175 CACTGGGGCCATCGAGCGACACGGAAGCCACAGAGGAGAGTGTCCCATGTGCATGACT 1234
QY 341 ValGlyProTyrValProLeuTyrPheGlyCysGluAlaGlyPheTyrValAspAlaGly 360
Db 1235 GTGGGGCCCTACCTCTCTCTGTGGGCTGTGAGGACAGATTTTATGTGCAANNNNNN 1294
QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlyIleSerAlaIleTyr 380
Db 1295 NNNNNNNCTATGTTTTCACCCCTGGGCGACGCTGTTCAGAGAGTCTGCCAAGTAC 1354
QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
Db 1355 TGGTTCGAGATCCCATCGCCCAAGACCAAGCCGTTTCAAGCCGCTCTCCGTTCTGC 1414
QY 401 AlaThrGlnLeuValGlyGlnIleAsnCysIleLeuLeuIlePheGlnIleProIleAsp 420
Db 1415 GCCACGACGCTGGTGTGTGAACAGAACTGCATCAATTTGATTTCCAAAGTCCAGTGAC 1474

```

```

FILE REFERENCE: 529452000121
CURRENT APPLICATION NUMBER: US/10/085,117
CURRENT FILING DATE: 2002-02-27
PRIOR APPLICATION NUMBER: US 09/798,586
PRIOR FILING DATE: 2001-03-02
NUMBER OF SEQ ID NOS: 361
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 284
LENGTH: 1734
TYPE: DNA
ORGANISM: Mus musculus
FEATURES:
NAME/KEY: variation
LOCATION: (1)..(1734)
OTHER INFORMATION: n = any nucleotide
US-10-085-117-284

Alignment Scores:
Pred. No.: 6.5e-234 Length: 1734
Score: 211.50 Matches: 388
Percent Similarity: 95.2% Conservative: 12
Best Local Similarity: 92.4% Mismatches: 19
Query Match: 92.2% Indels: 1
DB: Gaps: 1

US-10-041-030-4 (1-420) x US-10-085-117-284 (1-1734)
QY 1 MetPheSerProGlyGlnGlnGluHisCysAlaProAsnGluProValIleTyrGly 20
Db 218 ATGTTTCCCGGGCGAGAGGAAACCAAGCGCCCAACAGAGACCGGTGAATACAGG 277
QY 21 GlnLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyAspArgGly 40
Db 278 GACCTGTGTCTCTGGGAGTACATGTGTCTTAACTTAATGAGACAGGGGCAAGAAAGAA 337
QY 41 SerArgPheAlaLeuTyrIleArgProValAlaAsnGlyValIlePheProSerThrValHis 60
Db 338 AGCAATTTACCTCTTATTAAGGAGACTACGACGATGTGTCAAAACCGACACATCCAC 397
QY 61 ValIleSerThrProGlnAlaSerIleValIleSerCysIleGlyGlnHisSerIleSer 80
Db 398 ATGTTTCCACACACACAGGGGTCCAAAGGCATCAGCTCCAGAGACATCAACATATCG 457
QY 81 TyrThrLeuSerArgAsnGlnThrValValGluTyrThrHisAspIleAspThrAsp 100
Db 458 TACACGTTGTCACGCGGACGAGCGGTAGTGTGAGATACACACACATTAAGACACAGAC 517
QY 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValThrAspThrIle 120
Db 518 ATGTTTCAGGTGGGCGAGTCAACAGAAAGCCCACTTACCTTCGGTGTACAGACCGGTT 577
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 578 TCCGGGGGCGTCAAGAACGAAGAT---GCCCATGATCAACAGACGACCATCTCTAGGTTGCA 634
QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
Db 635 TGCAGAGTCTGTGTGACAGAAACGACCATACACAGCAGCATATTCGGCGACGAGATTC 694
QY 161 AspSerSerIleValIlePheLeuGlyGluValAlaIleValIlePheIleAspProAspGly 180
Db 695 GATTCTTCCAAAATATCTTTCTTGGAGAAAGACGAAATGAAATGAAATGAAATGAAATGAA 754
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
Db 755 CACATGATGATGACTCACTACCAACGGTGTCTGTGTATGACCCGCAAGAGGCTTACCC 814
QY 201 GlnGlnSerGlnProGlyValITPArgGluIleSerValCysGlyAspValITyrThrLeu 220
Db 815 GAGGAATCCAGCCTGAGTCTGAGAGAGATCTGTCTGTGGGAGTATATACACTTG 874
QY 221 ArgGluThrArgSerArgAsnGlnArgGlyValLeuValGlnSerGluThrAsnValLeu 240

```

```

RESULT 13
US-10-085-117-284
; Sequence 284, Application US/10085117
; Publication No. US20030232334A1
; GENERAL INFORMATION:
; APPLICANT: Morrish, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR CANCER

```

```

Db      875 CGAGAGACGAGTCCGCCAGACAGAGGAGAAAGCTGGTGAAGATGAGACCAACGTCCTG 934
Qy      241 GAAAGGYSerleuileapleuCyeglyAlaThrleuLeuTPAaGThraIaapGly 260
Db      935 CAAGAGGCTCCTCATTAACCTGTGTGGGACCACTTCTCTGGAGAACCGAGATGGC 994
Qy      261 LeuPheHleThrProThrglnlyshIleGluAlaLeuArglnGlnIuIleAmaIaAla 280
Db      995 CTTTTCACGCTCCTACTACAGAACACATGAAAGCCCTCCGGACGAGATCAATGACGCC 1054
Qy      281 ArgProGlnCyseProValGlyLeuAenThrLeuAlaPheProSerIleAmaArglyGlu 300
Db      1055 CGACCCAGAGGCCCCGGGGCCCTTAAACACCCCTGGCTTCCCGACATCAACCGAAGGAA 1114
Qy      301 ValValGluGluArgGlnProTPAlaTyrlauserCyeglyshIleValIleGlyThrHls 320
Db      1115 GTGGTGAAGAGAACGAGCCCTGGGCACTGAGCTGAGCTGGCCAGTGGACCGGTTACAC 1174
Qy      321 AenTPGlyHlsArgSerAepThrgluAlaAenGluArglyGlyProMetCyseArgThr 340
Db      1175 CACTGGGGCCATCGAGCGACACGGAAGCCACGAGAGGAGATGTCCCATGTGATGACT 1234
Qy      341 ValGlyProTyrlaProleuTPLeuGlyCyeglyAlaGlyPheTyrlaAapAlaGly 360
Db      1235 GTGGGCCCCCTACGCTCTCTGTGGCTGGCTGAGGACAGATTTTATGTGANNNNNN 1294
Qy      361 ProProThraIaPheThrProCyeglyHlsValCyseSerGlyuysSerAlaIleTyrla 380
Db      1295 NNNNNNNCTCATGTTTTCACCCCTGGGCACTGTGTCAGAGAACTGGCCAGATAC 1354
Qy      381 TrpSerGlnIleProleuProHlsGlyThrHlsAlaPheHlsAlaAlaCyseProPheCys 400
Db      1355 TGGTGGAGATCCCACTGCCCCACGGAAGCAGCGTTTCAAGCCGCTGTCCGTTCTGC 1414
Qy      401 AlaThrGlnLeuValGlyGluGlnAenCyseIleValleuIlePheGlnGlyProIleAap 420
Db      1415 GCCACGACGCTGGTGTGTGAACAGAACTGCATCAATTTGATTTCCAGGTTCAGTGGAC 1474

```

RESULT 14

```

US-09-843-905A-3
; Sequence 3, Application US/09843905A
; Patent No. US20020168683A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843, 905A
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-843-905A-3

```

Alignment Scores:

```

Pred. No.: 1,49e-211 Length: 1257
Score: 1917.00 Matches: 342
Percent Similarity: 90.5% Conservative: 38
Best Local Similarity: 81.4% Mismatches: 38
Query Match: 83.7% Indels: 2
Gaps: 1

```

US-10-041-030-4 (1-420) x US-09-843-905A-3 (1-1257)

```

Qy      1 MecPheSerProGlnGlnGluHlsCyseAlaProAenlyGluProValIleTyrlaGly 20
Db      1 ATGTTTCTCTCTGATCAAGAAATCAT-----CCATCTAAAGCACCACTAAATATATGCT 54

```

```

Qy      21 GluLeuValIleGlnGlyTyrlaenGlyAlaLeuProAenGlyAapArglyAarglyGly 40
Db      55 GAACCTATTGCTTAGAGATATATATGATCTCTCCAAACGGTATAGAGAAAGAGAAA 114
Qy      41 SerArgPheAlaLeuTyrlaGlyAargProlyAlaAenGlyValIleAenProSerThraIleHls 60
Db      115 AGTAGGTTGGCTTGTATTAACCACTTAAGCAAAAGGGGTGAAGCCAGCATGTGCAT 174
Qy      61 ValIleSerThraIleAenThraIleSerCyseIleGlyGlnHlsSerIleSer 80
Db      175 ATTGCTTGTACTCTCAGGCTGCAAGGCAATAGCAACAAACAGCATAGCATATCA 234
Qy      81 TyrlaThraSerArgAenThraValIleValGlyTyrlaThraIleAapTyrlaAapThra 100
Db      235 TATACCTTATCTGGGGCCAGACTGTGTGGTTGAATATATCTATACAGCAACCGAT 254
Qy      101 MecPheGlnValGlyArgSerThrglyuserProIleAapPheValIleThraPThrIle 120
Db      295 ATGTTTCAGATTTGGCCGCTGACCTGAAGCCCATTTGATTTGATTACTGACAGGCTT 354
Qy      121 SerGlySerGlnAenThraPheGlnIleThrglnSerThraIleSerArgPheAla 140
Db      355 CTGGAAGTCAAGTAATCTGATACACAGTCACTAACAGCATATACAGATTTTGGC 414
Qy      141 CyseArgIleValCyseAapArgAenGluProTyrlaThraIleAargIlePheAlaAlaGlyPhe 160
Db      415 TGCAGATATCATATGTAACGGAATCCTCCTTACAGCAGGATTTATGTGTGACGGGTTT 474
Qy      161 AapSerSerIleAenIlePheLeuGlyGluValAlaIleTyrlaIleAapProAapGly 180
Db      475 GACTCATCAAAAAACATCTTTCTGGGAGAGAGGCTGCCMAAGGAAGCATGACATGAGA 534
Qy      181 HlsMecAapGlyLeuThraThraAenGlyValIleValMecHlsAproArglyGlyPheThr 200
Db      535 CAGATGATGGCTTGAACCATATAGGTTCTGTATGATCATCAGCAATGGGTTTACA 554
Qy      201 GluGluSerGlnProGlyValIleTPAArgGluIleSerValCyseGlyAapValTyrlaLeu 220
Db      595 GAAGACTCCAGCCTGGAATATGAGAAATATCGGTGTGGAAATGATTTAGCCTTA 654
Qy      221 ArgGluThraArgSerAlaGlnGlnAarglyGlyLeuValGluSerGluThraAenValLeu 240
Db      655 CGTGAACCAAGATCGGCTCAGCAGAGAGAAATGTGGAAATTTGAACCAATCAGTTA 714
Qy      241 GlnAapGlySerleuileapleuCyeglyAlaThrleuLeuTPAaGThraIaapGly 260
Db      715 CAAGATGGCTCGTTAATGACCTGTGTGCAACATTTGATGTGGTATGCAAGAGGC 774
Qy      261 LeuPheHleThrProThrglnlyshIleGluAlaLeuArglnGlnIuIleAmaIaAla 280
Db      775 CTTTCCACACTCTCACCAGAAACATTTAAGAGCTTTAAGACAGAAATCAATGACGCA 834
Qy      281 ArgProGlnCyseProValGlyLeuAenThrLeuAlaPheProSerIleAmaArglyGlu 300
Db      835 CGACCTCAGGCGCTGTAGGGTTCAACACACTGACTTCTAGTATGAAGAGAAAGAC 894
Qy      301 ValValGluGluArgGlnProTPAlaTyrlauserCyeglyshIleValIleGlyThrHls 320
Db      895 GTGTGATGATAAAACAAACCATGGGTATCTTAACGCGGCAATGATACATGGCTATCAT 954
Qy      321 AenTPGlyHlsArgSerAepThrgluAlaAenGluArglyGlyProMetCyseArgThr 340
Db      955 AACTGGGAAACAAAGAAAGCTGATGGCAAGATCGTGAATGTCTATGTGTAGTCT 1014
Qy      341 ValGlyProTyrlaProleuTPLeuGlyCyeglyAlaGlyPheTyrlaAapAlaGly 360
Db      1015 GTTGGTCCCTATGTTCTCTGTGGCTTGAATGAGCTGATTTATGTCAGCGCGGC 1074
Qy      361 ProProThraIaPheThrProCyeglyHlsValCyseSerGlyuysSerAlaIleTyrla 380
Db      1075 CTTCAACCATCGTTTAAAGCCGTGTGGCAATGTGTGTCAGAAAGAACATGCTCAT 1134
Qy      381 TrpSerGlnIleProleuProHlsGlyThrHlsAlaPheHlsAlaAlaCyseProPheCys 400

```


Db 1135 TGGTCCAGATCCCACTTCTCATGTACTCATCTTTTCATGACGCTGCTCCCTTTGT 1194
 Qy 401 AlAThGlnLeuValGlyGlnAenCysIleLeuLeuIlePheGlnGlyProIleAsp 420
 Db 1195 GCACATCAGTTGGCTGGTGAACAAGGCTACATCAGACTTATTTTTCAGAGCCTTAGAC 1254

RESULT 15

US-10-317-250-3
 ; Sequence 3, Application US/10317250
 ; Publication No. US20030165945A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bird, Timothy A.
 ; APPLICANT: Cosman, David J.
 ; APPLICANT: Li, Xiaoxia
 ; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
 ; FILE REFERENCE: 2990-B
 ; CURRENT APPLICATION NUMBER: US/10/317,250
 ; CURRENT FILING DATE: 2002-12-11
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 3
 ; LENGTH: 1257
 ; TYPE: DNA
 ; ORGANISM: Homo sapiens
 US-10-317-250-3

Alignment Scores:

Pred. No.: 1,49e-211 Length: 1257
 Score: 1917.00 Matches: 342
 Percent Similarity: 90.5% Conservative: 38
 Best Local Similarity: 81.4% Mismatches: 38
 Query Match: 83.7% Indels: 2
 Gaps: 1

US-10-041-030-4 (1-420) x US-10-317-250-3 (1-1257)

Qy 1 MetPheSerProGlyGlnGluHisCysAlaProAsnLysGluProValLysTyrGly 20
 Db 1 ATGTTTCTCTCATCAAGAAATCAT-----CCATCTAAGCACAGTAAATATAGGT 54
 Qy 21 GlnLeuValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgGly 40
 Db 55 GAACTCATTTGTCTTGAATTAATGATCTCTCCAAACGGTATGAGAGAGAGAAA 114
 Qy 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
 Db 115 AGTAGTTTCTTGTTTAAGAAGCTAAGCAAAATGGGTGAAGCCACACTGTGCAT 174
 Qy 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80
 Db 175 ATTGCTTGTATCTCTCAGGCTGCAAGGCAATTAAGCAACAAGCACAGTACATATCA 234
 Qy 81 TyrThrLeuSerArgAsnGlnThrValValGluTyrThrHisAspLysAspThrAsp 100
 Db 235 TATACCTTATCTGGGCCCAAGACTGTGTGATGATATCTCATGACCAACACCGAT 294
 Qy 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValValThrAspThrIle 120
 Db 295 ATGTTTCAGATTGGCCCGCTGACCTGAAGAGCCCATTTTGTAGTACTGACACGGTT 354
 Qy 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 355 CCTGGAAGTCAAAATTTCTGATACACAGTCAAGTCAAAAGCACTATATCAAGATTGCC 414
 Qy 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
 Db 415 TGCAGAAATCATATGTGAACGAATCTCTCTTACAGACGAGATTATGCTGCAGGGTTT 474
 Qy 161 AspSerSerLysAsnIlePheLeuGlyGluValAlaIleValTyrLysAsnProAspGly 180
 Db 475 GACTCATCAAAAACATCTTTCTTGGGGAAGAGGCTGCCAAATGGAAGACATCAGATGA 534

Qy 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
 Db 535 CAGATGATGGCTTGAACCACTAATGTGTTCTTGTCAATGACATCCAGCAATGGTTTACA 594
 Qy 201 GlnGluSerGlnProGlyValIleTyrArgGlnIleSerValCysGlyAspValTyrThrLeu 220
 Db 595 GAAGACTCCAGGCTGGAAATGAGAGAAATATCGGTGTGTGGAAATGATATTACCTTA 654
 Qy 221 ArgGluThrArgSerAlaGlnGlnArgGlyValLeuValGluSerGlnThrAsnValLeu 240
 Db 655 CCGGAACCAAGATCGGCTCAGACAGAGAGAAAATGTGAAATTTGAACCAATCAGTTA 714
 Qy 241 GlnAspGlySerLeuIleAspLeuCysGlyValAlaThrLeuThrProArgThrAlaAspGly 260
 Db 715 CAAAGATGCTCGTTAATTAATGACCTGTGTGGTGAACATTTTATGGCTTACGCAAGGC 774
 Qy 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGlnIleAsnAlaAla 280
 Db 775 CTTTCCACACTCTTACCGTGAAGCATTTAGAAGCTTTAAGACAGGAATCAATGACAGA 834
 Qy 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
 Db 835 CGACCTCAGTGGCCCTGTAGGGTTTCAACACATGCACTTCTTCAATGAAAGGAAAGAC 894
 Qy 301 ValValGluGluLysGlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
 Db 895 GTTGATGATGAAGAAACACACATGGGTATATCTTAACCTGGGCCCATGTACATGCTATCAT 954
 Qy 321 AsnTyrGlyHisArgSerAspThrGlnAlaAsnGluArgGluCysProMetCysArgThr 340
 Db 1075 CCTCAACCCATGCGTTTAGCCCGTGTGGCAGATGTGTTCGAAAAGACAACTGCTAT 1134
 Qy 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
 Db 1135 TGGTCCAGATCCCACTTCTCATGTACTCATCTTTTCATGACACGCTGCTTTTGT 1194
 Qy 401 AlAThGlnLeuValGlyGlnAenCysIleLeuLeuIlePheGlnGlyProIleAsp 420
 Db 1195 GCACATCAGTTGGCTGGTGAACAAGGCTACATCAGACTTATTTTTCAGAGCCTTAGAC 1254

Search completed: March 2, 2006, 07:20:48
 Job time : 1092 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 2, 2006, 05:56:29 ; Search time 253 Seconds
(without alignments)
2950.893 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MESPGEHNCAPKPEPVKVG.....ATQVGEQNCIKIFQGPID 420

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 1303057 seqs, 888780828 residues

Total number of hits satisfying chosen parameters: 2606114

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-Q=/abses/ABSSWEB.spool/US10041030/runat_01032006_134408_21912/app_query.fasta_1
-DB=Issued Patents NA -QFMT=fasta -SUFFIX=p2n.rn1 -MINMATCH=0 -LOOPTCL=0
-LOOPTCL=0 -UNITS=bits -START=1 -END=1 -MATRIX=blosum62 -TRANS=human0.cdt
-LIST=45 -DOCALL=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15
-MODE=LOCAL -OUTWT=pct -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000
-HOST=abses04 -USER=US10041030 @CCN 1 1 193 @runat_01032006_134408_21912
-NCPUS=6 -ICPU=3 -NO MMAP -NEG SCORES=0 -WAIT -DSPBLOCK=100 -LONGLOG
-DEV TIMEOUT=120 -WARM TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6
-FGAPEXT=7 -YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database : Issued Patents NA:
1: /cgn2_6/prodata/1/ina/1 COMB.seq:*
2: /cgn2_6/prodata/1/ina/5 COMB.seq:*
3: /cgn2_6/prodata/1/ina/6A COMB.seq:*
4: /cgn2_6/prodata/1/ina/6B COMB.seq:*
5: /cgn2_6/prodata/1/ina/H COMB.seq:*
6: /cgn2_6/prodata/1/ina/PCUS COMB.seq:*
7: /cgn2_6/prodata/1/ina/PP COMB.seq:*
8: /cgn2_6/prodata/1/ina/RE COMB.seq:*
9: /cgn2_6/prodata/1/ina/backfile1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|-------|--------------------|
| 1 | 2283 | 99.7 | 1263 | 3 | US-09-843-905A-7 |
| 2 | 2175.5 | 95.0 | 1260 | 3 | US-09-843-905A-5 |
| 3 | 1917 | 83.7 | 1257 | 3 | US-09-843-905A-3 |
| 4 | 1912 | 83.5 | 1257 | 3 | US-09-843-905A-1 |
| 5 | 1642 | 71.7 | 1338 | 3 | US-09-843-905A-11 |
| 6 | 1482.5 | 64.7 | 2508 | 3 | US-10-104-047-1958 |
| 7 | 120.5 | 5.3 | 2750 | 2 | US-08-479-403-1 |
| 8 | 120.5 | 5.3 | 2750 | 2 | US-08-479-403-1 |
| 9 | 120.5 | 5.3 | 2750 | 3 | US-08-835-734-1 |

| | | | | | | |
|------|-------|-----|---------|---|----------------------|-------------------|
| C 10 | 113.5 | 5.0 | 4403765 | 3 | US-09-103-840A-2 | Sequence 2, Appl1 |
| C 11 | 113.5 | 5.0 | 4411529 | 3 | US-09-103-840A-1 | Sequence 1, Appl1 |
| C 12 | 109 | 4.8 | 1580 | 3 | US-09-252-991A-14397 | Sequence 14397, A |
| C 13 | 109 | 4.8 | 3801 | 3 | US-09-252-991A-14280 | Sequence 14280, A |
| C 14 | 107 | 4.7 | 86414 | 3 | US-09-949-016-12345 | Sequence 12345, A |
| C 15 | 107 | 4.7 | 86414 | 3 | US-09-949-016-12345 | Sequence 12345, A |
| C 16 | 106.5 | 4.7 | 2240 | 3 | US-09-799-451-571 | Sequence 571, A |
| C 17 | 106 | 4.6 | 1314 | 3 | US-09-902-540-2044 | Sequence 2044, A |
| C 18 | 106 | 4.6 | 2574 | 3 | US-09-902-540-4745 | Sequence 4745, A |
| C 19 | 106 | 4.6 | 25254 | 3 | US-09-902-540-1233 | Sequence 1233, A |
| C 20 | 106 | 4.6 | 4403765 | 3 | US-09-103-840A-1 | Sequence 1, Appl1 |
| C 21 | 106 | 4.6 | 4411529 | 3 | US-09-103-840A-2 | Sequence 2, Appl1 |
| C 22 | 105 | 4.6 | 6476 | 3 | US-09-854-133-428 | Sequence 428, A |
| C 23 | 105 | 4.6 | 22512 | 3 | US-09-902-540-1220 | Sequence 1220, A |
| C 24 | 105 | 4.6 | 34030 | 3 | US-09-949-016-12248 | Sequence 12248, A |
| C 25 | 105 | 4.6 | 34031 | 3 | US-09-949-016-13540 | Sequence 13540, A |
| C 26 | 104 | 4.5 | 2133 | 3 | US-09-252-991A-9129 | Sequence 9129, A |
| C 27 | 104 | 4.5 | 44120 | 3 | US-09-949-016-14151 | Sequence 14151, A |
| C 28 | 104 | 4.5 | 44120 | 3 | US-09-949-016-14152 | Sequence 14152, A |
| C 29 | 104 | 4.5 | 44120 | 3 | US-09-949-016-14153 | Sequence 14153, A |
| C 30 | 104 | 4.5 | 44120 | 3 | US-09-949-016-14154 | Sequence 14154, A |
| C 31 | 104 | 4.5 | 44120 | 3 | US-09-949-016-14155 | Sequence 14155, A |
| C 32 | 104 | 4.5 | 44120 | 3 | US-09-949-016-14156 | Sequence 14156, A |
| C 33 | 103.5 | 4.5 | 2670 | 3 | US-09-902-540-517 | Sequence 517, A |
| C 34 | 103 | 4.5 | 30001 | 2 | US-08-125-468-1 | Sequence 1, Appl1 |
| C 35 | 103 | 4.5 | 30001 | 2 | US-08-474-933-1 | Sequence 1, Appl1 |
| C 36 | 102.5 | 4.5 | 4695 | 3 | US-09-902-540-8855 | Sequence 8855, A |
| C 37 | 102.5 | 4.5 | 3059 | 3 | US-09-902-540-943 | Sequence 943, A |
| C 38 | 101.5 | 4.4 | 53615 | 3 | US-09-902-540-540 | Sequence 540, A |
| C 39 | 101.5 | 4.4 | 234884 | 3 | US-09-214-808-1 | Sequence 1, Appl1 |
| C 40 | 100.5 | 4.4 | 4041 | 3 | US-09-949-016-16420 | Sequence 16420, A |
| C 41 | 100 | 4.4 | 4041 | 3 | US-09-949-016-1305 | Sequence 1305, A |
| C 42 | 100 | 4.4 | 4041 | 3 | US-09-949-016-1306 | Sequence 1306, A |
| C 43 | 100 | 4.4 | 4416 | 3 | US-09-919-039-116 | Sequence 116, A |
| C 44 | 100 | 4.4 | 8930 | 3 | US-09-077-098A-1 | Sequence 1, Appl1 |
| C 45 | 100 | 4.4 | 8930 | 3 | US-10-192-584-1 | Sequence 1, Appl1 |

ALIGNMENTS

RESULT 1
US-09-843-905A-7
Sequence 7, Application US/09843905A
Patent No. 6703487
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
INVENTOR: Bird, Timothy A.
TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
FILE REFERENCE: 2990-A
CURRENT APPLICATION NUMBER: US/09/843, 905A
PRIORITY FILING DATE: 2001-04-27
PRIORITY APPLICATION NUMBER: US 60/200,198
PRIORITY FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patent version 3.1
SEQ ID NO 7
LENGTH: 1263
TYPE: DNA
ORGANISM: Homo sapiens
US-09-843-905A-7
ALIGNMENT Scores:
Pred. No.: 1,38e-238
Score: 2283.00
Percent Similarity: 99.8%
Best Local Similarity: 99.8%
Query Match: 99.7%
DB: 3
Gaps: 0
US-10-041-030-4 (1-420) x US-09-843-905A-7 (1-1263)
1 MetaphaseProGlycylngluGluHisCysAlaProAsnIleGluProValIleTyrGly 20
|||||

Db 1 ATGTTTCCCTGCGCAGAGAACTGCGCCCAATAAGAGCCAGTGAATAACGGG 60
 Qy 21 GlnLeuValValLeuGlyTyrAenGlyValAleuProAenGlyAAspArgIyAArgIy 40
 Db 61 GAGCTGGTGGTCTCGGGTCAATGGTGTCTTACCCCAATGAGATGAGAGCGAGGAAA 120
 Qy 41 SerArgPheAlaLeuTyrIyAArgProLyAAlaAenGlyValIyAProSerThrValHis 60
 Db 121 AGTAGATTTGCCCTCTACAGACGGGCCCAAGCAATGTGTCAAAACCCAGACGCTCAT 180
 Qy 61 ValIleSerThrProGlnAlaSerIyAAlaIleSerCyAlySglYglnHisSerIleSer 80
 Db 181 GTGATATCCACGGCCCGCAGCATCCAGGCTATCAGCTGCAAAAGTCAACAGATATATCC 240
 Qy 81 TyrThrLeuSerArgAenGlnThrValValIgluTyrThrHisAAspIyAAspThrAsp 100
 Db 241 TACACTTTCAGAGAAATCAGACTGTGTGTGTGATGACACATGATTAAGATACGGAT 300
 Qy 101 MetPheGlnValIgluArgSerThrGluSerProIleAAspPheValIThrAAspThrIle 120
 Db 301 ATGTTTCAGGTGGGAGATCAACAGAAAGCCCTATGCACTTCGTTCACAGACAGATTC 360
 Qy 121 SerGlySerGlnAenThrAAspGluAglIleThrGlnSerThrIleSerArgPheAla 140
 Db 361 TCTGGCAGCCAGAACAGGAGCAGAACCCAGATTCACAGAGACACATATCCAGGTTCC 420
 Qy 141 CyAArgIleValCyAAspArgAenGluProTyrThrIleAArgIlePheAlaAglIyPhe 160
 Db 421 TGCAGAGTCGTGTGCGACAGAAATGAACCTTACACAGCAGAGATTCGCGCGGATTT 480
 Qy 161 AspSerSerIyAAsnIlePheLeuGlyGluValAlaIleValIleValIleValIleVal 180
 Db 481 GACTCTTCCAAAACATATTTCTTGAGTAAAGCACAAGAGTGAAGAAACCCGACGCGC 540
 Qy 181 HisMetAAspGlyLeuThrThrAenGlyValIleuValIleHisAAspArgIyGlyPheThr 200
 Db 541 CACATGATGGGCTCACTACTAATGAGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
 Qy 201 GlnGlnSerGlnProGlyValIleTPAArgIleIleSerValCyAglYAspValIleThrLeu 220
 Db 601 GAGGAGTCCAGCCCGGGGTCTGCGCGAGATCTCTGTGTGTGTGTGTGTGTGTGTGTGTGT 660
 Qy 221 ArgGlnThrArgSerAlaGlnGlnArgIyIyLeuValIgluSerGlnThrAAsnValIleu 240
 Db 661 CGAGAAACAGGTGCGGCCAGCAACAGAAAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
 Qy 241 GlnAAspGlySerLeuIleAAspLeuCyAglYAlaThrLeuLeuTyrParGlyThrIleAAspGly 260
 Db 721 CAGGACGGCTCCCTCATTTGACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
 Qy 261 LeuPheHisThrProThrGlnIyHisIleIleGluAlaLeuAArgGlnGluIleAAsnAla 280
 Db 781 CTTTTCATATCTCCACATCAGAACACATAGAACCTTCGCGCAGAGATTAAGCGCGCC 840
 Qy 281 ArgProGlnCyAProValIgluAAsnThrIleuAlaPheProSerIleAAsnArgIyGlu 300
 Db 841 CGGCTCAGGTGCTGTGGGCTTACACCTGCTCTCCAGACATCAACAGAAAGAG 900
 Qy 301 ValValIgluGluIyGlnProTyrAlaTyrIleuSerCyAglYHisValHisIleGlyTyrHis 320
 Db 901 GTGGTGAAG 960
 Qy 321 AAsnTyrGlyHisAAspSerAAspThrGluAlaAenGluAArgIyCyAProMetCyAAspThr 340
 Db 961 AACTGGGGCCCATTCGAGTGAACCGAGAGCCCAACAGAGAGAGAGTGTCCCATGTGACAGACT 1020
 Qy 341 ValIgluProTyrValProLeuTyrLeuGlyCyAglIleAglIyPheTyrValAAspAlaGly 360
 Db 1021 GTGGGCCCCCTATGTGCTCTCTGTGGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
 Qy 361 ProProThrHisAlaPheThrProCyAglYHisValCyAserGlnIyAAsnValIleTyr 380
 Db 1081 CCGGCAACTCATGTCTTCACTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140

Qy 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCyAProPheCyA 400
 Db 1141 TGGTTCAGATTCGCCGTGTGCTCATGGAACATGATTCATGACGCTGTGCTGTGCTGTGCT 1200
 Qy 401 AlaThrGlnLeuValIgluGlnAenCyAglIleValIlePheGlnIyProIleAAsp 420
 Db 1201 GTTACACAGCTGT 1260
 RESULT 2
 US-09-843-905A-5
 ; Sequence 5, Application US/09843905A
 ; Patent No. 6703487
 ; GENERAL INFORMATION:
 ; APPLICANT: Bird, Timothy A.
 ; APPLICANT: Cosman, David J.
 ; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
 ; FILE REFERENCE: 2990-A
 ; CURRENT APPLICATION NUMBER: US/09/843, 905A
 ; CURRENT FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: US 60/200,198
 ; PRIOR FILING DATE: 2000-04-28
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: PatentIn version 3.1
 ; SEQ ID NO 5
 ; LENGTH: 1260
 ; TYPE: DNA
 ; ORGANISM: Mus musculus
 US-09-843-905A-5
 Alignment Scores:
 Pred. No.: 7,05e-227 Length: 1260
 Score: 2175,50 Matches: 399
 Percent Similarity: 97.4% Conservative: 10
 Best Local Similarity: 95.0% Mismatches: 10
 Query Match: 95.0% Indels: 1
 Gaps: 1
 US-10-041-030-4 (1-420) x US-09-843-905A-5 (1-1260)
 Qy 1 MetPheSerProGlyGlnGlnGluHisCyAAlaProAenGluProValIySgIy 20
 Db 1 ATGTTTCCCGGGCCAGAGAAACCCAGCGCCCAAGAGAGCGGTGAATAACGGG 60
 Qy 21 GlnLeuValValLeuGlyTyrAenGlyValAleuProAenGlyAAspArgIyAArgIy 40
 Db 61 GAGCTGGTGGTCTCGGGTCAATGGTGTCTTACCCCAATGAGATGAGAGCGAGGAAA 120
 Qy 41 SerArgPheAlaLeuTyrIyAArgProLyAAlaAenGlyValIyAProSerThrValHis 60
 Db 121 AGTAGATTTGCCCTCTACAGACGGGCCCAAGCAATGTGTCAAAACCCAGACGCTCAT 180
 Qy 61 ValIleSerThrProGlnAlaSerIyAAlaIleSerCyAlySglYglnHisSerIleSer 80
 Db 181 GTGATATCCACGGCCCGCAGCATCCAGGCTATCAGCTGCAAAAGTCAACAGATATATCC 240
 Qy 81 TyrThrLeuSerArgAenGlnThrValValIgluTyrThrHisAAspIyAAspThrAsp 100
 Db 241 TACACTTTCAGAGAAATCAGACTGTGTGTGTGATGACACATGATTAAGATACGGAT 300
 Qy 101 MetPheGlnValIgluArgSerThrGluSerProIleAAspPheValIThrAAspThrIle 120
 Db 301 ATGTTTCAGGTGGGAGATCAACAGAAAGCCCTATGCACTTCGTTCACAGACAGATTC 360
 Qy 121 SerGlySerGlnAenThrAAspGluAglIleThrGlnSerThrIleSerArgPheAla 140
 Db 361 TCGGGCGGTGAGAACGAAAGT---GCCAGATCACAAGAGCACCATCTCTAGTTCGCA 417
 Qy 141 CyAArgIleValCyAAspArgAenGluProTyrThrIleAArgIlePheAlaAglIyPhe 160
 Db 418 TGCAGAGTCGT 477
 Qy 161 AspSerSerIyAAsnIlePheLeuGlyGluValAlaIleValIleValIleValIleVal 180

```

|||||
Db 478 GATTCTTCCAAAATATCTTCTTGAGAGAAAGCAAAATGSAAAAACCTGATGGA 537
Qy 181 HiMeAspGlyLeuThrThrAsnGlyValLeuValMeHiAspProArgGlyGlyPheThr 200
Db 538 CACATGATGATGACTACTACCAATGGTGCTCTAGTATGACCCGCAAGAGGCTTCACC 597
Qy 201 GluGluSerGlnProGlyValTTPArgGluIleSerValCysGlyAspValTyrThrLeu 220
Db 598 GAGGAATCCAGAGCTGAGTGTGAGAGAAATCTCTGTGTGGGAGTGTACACCTTG 657
Qy 221 ArgGluThrArgSerAlaGlnGlnArgGlyValLeuValGluSerGluThrAsnValLeu 240
Db 658 CGAGAGACCAAGCTCGGCCGAGAGAGGGAAGCTGTGTGAAGTGAACCAACCTCTG 717
Qy 241 GluAspGlySerLeuIleAspLeuCysGlyValAlaThrLeuLeuTTPArgThrAlaAspGly 260
Db 718 CAAAGCGGCTCCCTCATGTGACCTGTGTGGGCACTCTCTCTGAGAAACCGCAGATGGC 777
Qy 261 LeuPheHisThrProThrGlnGlyHisIleGluAlaLeuArgGlnIleAsnAlaAla 280
Db 778 CTTTTCAGGCTCTTACTACAGAGCAATGAAAGCCCTCCGCAAGATCAATGACGCC 837
Qy 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLeuGlu 300
Db 838 CGACCCCAAGTCCCGGCGGCTTAACACCTTGCCCTCCCAAGATCAACCGAAGAA 897
Qy 301 ValValGluGluValysGlnProTTPAlaTyrLeuSerCysGlyValHisGlyTyrHis 320
Db 898 GTGTGTGAAGAAGAGAGCCCTGTGGCATACCTGAGCTGCGGCATGTGACCGGCTACAC 957
Qy 321 AsnTTPGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
Db 958 AGCTGGGGCCATGAGAGCGAGCGGAAAGCCCAAGAGAGAGTGTCCCAATGTCCAGACT 1017
Qy 341 ValGlyProTyrValProLeuTTPLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
Db 1018 GTGGGCCCTCACTCTCTCTGTGGCTGTGAGCAGATTTTATGTGATGCGGGA 1077
Qy 361 ProProThrHisAlaPheThrProCysGlyValAlaCysSerGluYsSerAlaYsTyr 380
Db 1078 CCCCACACTCAGCTTTCACTCCCTGCGGCACTGTTCAGAAAGTCTGCAAGTAC 1137
Qy 381 TTPSerGlnIleProLeuProHisGlyTyrHisAlaPheHisAlaAlaCysProPheCys 400
Db 1138 TGTGTCCAGATCCCACTGCGCCCAAGCAAGCAGCGCTTTCATGCGCTCTCTGTC 1197
Qy 401 AlaThrGlnLeuValGlyGluGlnAsnCysAlaIleAsnLeuIlePheGlnGlyProIleAsp 420
Db 1198 GCCACGACGCTGGTGTGTGAAACAGAACTGCAATTAATTTCCAAAGTCCAGTGGAC 1257

RESULT 3
US-09-843-905A-3
; Sequence 3, Application US/09843905A
; Patent No. 6703487
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 3
; LENGTH: 1257
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-843-905A-3
Alignment Scores:

```

```

Pred. No.: 1,02e-198 Length: 1257
Score: 1917.00 Matches: 342
Percent Similarity: 90.5% Conservative: 38
Best Local Similarity: 81.4% Mismatches: 38
Query Match: 83.7% Indels: 2
DB: 3 Gaps: 1

US-10-041-030-4 (1-420) x US-09-843-905A-3 (1-1257)
Qy 1 MecPheSerProGlyGlnGluHisCysAlaProAsnGlyLeuProValTyrGly 20
Db 1 ATGTTTCTCTGATCAAGAAATCAT-----CCATCTAAACACACAGATAAATATGTGT 54
Qy 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyAlaArgGly 40
Db 55 GAACTATGTCTTAGAATTAATGAAATCTCTCCAAACGGTATGAGAGAGAGGAAA 114
Qy 41 SerArgPheAlaLeuTyrIleAspProValAlaAsnGlyValIleAspProSerThrValHis 60
Db 115 AGTAGGTTGCTTTGTTTAAAGACCTTAAGGCAAAATGGGTGAAGCCGACCTGTGCAT 174
Qy 61 ValIleSerThrProGlnAlaSerIleValIleSerCysAlaGlyGlnHisSerIleSer 80
Db 175 ATGCTTGTACTCTCTCAGGCTGCAAGGCAATAGCAACAAAGACGATACATATCA 234
Qy 81 TyrThrLeuSerArgAsnGlnThrValValGluTyrThrHisAspIleAspThrAsp 100
Db 235 TATACCTTATCTCTGGGCCCAAGCTGTGTGTGAATATCATATACACAGAACACCGAT 294
Qy 101 MetPheGluValGlyArgSerThrGluSerProIleAspPheValAlaThrAspThrIle 120
Db 295 ATGTTTCAGATTGACCGGTGCACTGAAGGCCATGATTTTGTAGTACTGACACGGTT 354
Qy 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 355 CCGGAAGTCAAAAGTATCTGTATACACAGTCAAGTCAAAAGCATATATCAAGATTGGC 414
Qy 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
Db 415 TCGAAGATCATATGTAAGCAAGAAATCCTCTTTCACGACGGAATTAATGTGACGGGTTT 474
Qy 161 AspSerSerIleAsnIlePheLeuGlyValValAlaAlaIleTTPAspProAspGly 180
Db 475 GACTCATCAAAAACATCTTCTTGGGAGAAAGCTGCAAAATGAAAGACATCAAGTGA 534
Qy 181 HiMeAspGlyLeuThrThrAsnGlyValLeuValMeHiAspProArgGlyGlyPheThr 200
Db 535 CAGATGATGATGCTTGAACACTAATGCTGTCTGTATGATGATCAAGCAATGGGTTACA 594
Qy 201 GluGluSerGlnProGlyValTTPArgGluIleSerValCysGlyAspValTyrThrLeu 220
Db 595 GAAAGCTCCAGGCTGGAATATGAGAAATATCGGTGTGTGGAATGATTAATTAAGCTTA 654
Qy 221 ArgGluThrArgSerAlaGlnGlnArgGlyValLeuValGluSerGluThrAsnValLeu 240
Db 655 CGTGAACCAAGAGCTGAGTCAAGAGAGAAATATGTGAATTAACCAATCACTTA 714
Qy 241 GluAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTTPArgThrAlaAspGly 260
Db 715 CAAAGTGGCTCGTTAATTAACCTGTGTGCAACATTTGTATAGGCTACTGCAAGAGC 774
Qy 261 LeuPheHisThrProThrGlnGlyHisIleGluAlaLeuArgGlnIleAsnAlaAla 280
Db 775 CTTTCCCACTCCTCAACCGGAAAGCATTTAAGAGCTTTAAGACAGAAATCAATGACGCA 834
Qy 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLeuGlu 300
Db 835 CGACCTCAGGCTCTTAGAGGTTCAACACATAGCAATTTCTTAGTATGAAGAGAAAGAC 894
Qy 301 ValValGluGluValysGlnProTTPAlaTyrLeuSerCysGlyValHisGlyTyrHis 320
Db 895 GTTGTAGTAAABAAACACATAGGTATATCTTAACCTGCGGCAATGATACATGCTATCAT 954

```

```

QY      321  AsnTrpGlyHisValSerSerAspThrCysIuLlaEngiUaTSGuLcYsProMetCysAtrpThr 340
Db      955  AACTGGGGAAACCAAGAAAGAAAGCTGATGGCAAAAGATCGTGAATGTCCTATGTTGTAAGTCT 1011
QY      341  ValGlyProTyrValProLeuTrpPheGlyCysGluValaglyPheTyrValAspAagly 360
Db      1015  GTTGTCCTCCTATGTTCTCTGTGGCTTGATGTGAAGCTGATTTTATGTGAGCCGGCC 1074
QY      361  ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaTyr 380
Db      1075  CCTCAACCCATGCGCTTGAAGCCCGTGGGCATGTGTCTCGAAGAAAGCAACTGCTAT 1133
QY      381  TrpSerGlnLysProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
Db      1135  TGGTCCCGAGTCCCACTTCTCTCATGTACTACTATTTCATGACAGCTGTCCCTTTGT 1199
QY      401  AlaTrpGlnLeuValGlyGluGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
Db      1195  GCACATCATGTTGGCGTGGTGAACCAAGCGTCAATCAAGACTTATTTTCAAGGACCTTAGAC 1255

RESULT 4
US-09-843-905A-1
/ Sequence 1, Application US/09843905A
/ Patent No. 6703487
/ GENERAL INFORMATION:
/ APPLICANT: Bird, Timothy A.
/ APPLICANT: Cosman, David J.
/ TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
/ FILE REFERENCE: 2990-A
/ CURRENT APPLICATION NUMBER: US/09/843,905A
/ PRIOR FILING DATE: 2001-04-27
/ PRIOR APPLICATION NUMBER: US 60/200,198
/ PRIOR FILING DATE: 2000-04-28
/ NUMBER OF SEQ ID NOS: 15
/ SOFTWARE: PatentIn version 3.1
/ SEQ ID NO 1
/ LENGTH: 1257
/ TYPE: DNA
/ ORGANISM: Mus musculus
US-09-843-905A-1

Alignment Scores:
Score: 3,576-198 Length: 1257
Percent Similarity: 1912.00 Matches: 341
Best Local Similarity: 90.5% Conservative: 39
Query Match: 81.2% Mismatches: 36
DB: 83.5% Indels: 2
Gaps: 3 1

US-10-041-030-4 (1-420) x US-09-843-905A-1 (1-1257)
QY      1  MetPheSerProGlyGlnGluGlnHisCysAlaProAsnLysGluProValLysTyrGly 20
Db      1  ATGTTTCTCCGAGTCAAGAAATCAT-----CCTTCAAGGCCCACTAATAATATGCG 54
QY      21  GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyValArgLys 40
Db      55  GAACCATGTCCTTGAATATATGATCTCTCCAAACGGTGAATAGAGGAGGAGAA 114
QY      41  SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
Db      115  AGTAGGTTGCTTGTGTTTAAAGACCTAAGCGCAATGGGGTAAAGCTTAGCACCGTGCAC 174
QY      61  ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80
Db      175  ATGTGATGTACTCTCCAGGCTCCCAAGGCATATAGCAACAGGACCAAGCATATGATATCA 234
QY      81  TyrThrLeuSerArgAsnGlnThrValValGlyTyrThrHisAspLysAspThrAsp 100
Db      235  TATATCTTATCTCGAGCCAGCGGTGGTGTGATATATCTCATGACAGCAACTGAT 294
QY      101  MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120

```

| | | | |
|---|------|--|------|
| Db | 295 | ATGTTTCAGATGGTTCGGTCAACTGAAAGTCCATATGATTTTGTATGTAATGACACCGTT | 354 |
| Qy | 121 | SeelYgeerGlnAenThrAaspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla | 140 |
| Db | 355 | CCTGGAAAGTCAGATGATATTCGGAACGACAGCAAGTACATCAAAAGCACTATATCAAGATTGGCC | 414 |
| Qy | 141 | CysArgIleValCysAraArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe | 160 |
| Db | 415 | TGTAGATCATATGTAGAGCGACAGTCCCTTTACAGCTCGGATTTATGCTGCAGGGTTT | 474 |
| Qy | 161 | AaspSerIlyAasnIlePheLeuGlyGluIlyAlaAlaIlySTPLyAaspProAaspGly | 180 |
| Db | 475 | GATTCATCAAAAACATCTTTCTTGGGAGAGAGGCTGCCAAGTGGAAAGACATCGATGGG | 534 |
| Qy | 181 | HisMetAaspGlyLeuThrThrAasnGlyValLeuValMetHisProArgGlyGlyPheThr | 200 |
| Db | 535 | CAGATGGAATGGCTTGAACCATTAAGAGATCTTGTATGATGCATCCAGTAAATGGGTTCACA | 594 |
| Qy | 201 | GluGluSerGlnProGlyValITPAArgGluIleSerValCysGlyAaspValITyThrIleu | 220 |
| Db | 595 | GAAACATCCAAACCTGGAAATATGAGAGAAATATCATATGTGGAAATGCTTTCAGTCTTG | 654 |
| Qy | 221 | ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAasnValIleu | 240 |
| Db | 655 | CGTAAACCGATGACGCTCAGCAGACAGAAAGATGGTGGAAATTTGAACCAATCAGCTA | 714 |
| Qy | 241 | GlnAaspGlySerLeuIleAaspLeuCysGlyValAThrLeuLeuITPAArgThrAlaAaspGly | 260 |
| Db | 715 | CAAGATGGCTCTTAATATGACCTTTGGTGGCAACTTTCGTGGCGTACATCGCAAGAGCG | 774 |
| Qy | 261 | LeuPheHisThrProThrGlnIlyHisIleGluAlaLeuAArgGlnIleAasnAlaAla | 280 |
| Db | 775 | CTTTCACATATCTCTACTCTGAGACACTTAACACTTTTAACACAGAGATCAATGCAAGCT | 834 |
| Qy | 281 | ArgProGlnCysProValGlyLeuAasnThrLeuAlaPheProSerIleAasnArgLysGlu | 300 |
| Db | 835 | CGCGCGCAGTGCCCTGTAGAGGTTCAACACATCAAGCTTCCCAAGTATGAAGAGAGAGAT | 894 |
| Qy | 301 | ValValGluGluIlyGlnProTProTAlaTyrLeuSerCysGlyHisValHisGlyTyrHis | 320 |
| Db | 895 | GTTGTATGATGAAGAAAGCAACCATGGGTATATCTTAATCTCGGCGCATATGCTCATGTATCAT | 954 |
| Qy | 321 | AsnITProGlyHisArgSerAaspThrGluAlaAsnGluIlyArgGluCysProMetCysArgThr | 340 |
| Db | 955 | AACGGGGAAACAAAGAAAGAAAGTGAAGGCAAAATGCTGAATGCTCATGTGTAAGTCT | 1014 |
| Qy | 341 | ValGlyProTyrValProLeuITPLeuGlyCysGluAlaGlyPheTyrValAaspAlaGly | 360 |
| Db | 1015 | GTTGTGCCATATGTCCTCTGTGGCTTGGAGTGAAGTGAATTTATGTGACGCGCGC | 1074 |
| Qy | 361 | ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr | 380 |
| Db | 1075 | CCTCCCAACCATGCTTAAAGCCCTGTGGGCACTGTGTTCAGAAAGAACACGCGCTTAC | 1134 |
| Qy | 381 | TyrSerGlnIleProLeuProHisGlyIThrHisAlaPheHisAlaAlaCysProPheCys | 400 |
| Db | 1135 | TGGTCCCAAGATCCCACTTCCATCAGTAAAGCACTTTCAATGACGCTGGCCCTTCGT | 1194 |
| Qy | 401 | AlaThrGlnLeuValGlyGluGlnAasnCysIleLysLeuIlePheGlnGlyProIleAasp | 420 |
| Db | 1195 | GCAATCATGTGGCTGGTGAACAAAGGCTATATCAAGACTTATTTTCCAAAGACCTTTAAGAC | 1254 |
| <p>RESULT 5 US-09-843-905A-11 ; Sequence 11, Application US/09843905A ; Patent No. 6703487 ; GENERAL INFORMATION: ; APPLICANT: Bird, Timothy A. ; APPLICANT: Cosman, David J. ; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES ; FILE REFERENCE: 2990-A ; CURRENT APPLICATION NUMBER: US/09/843, 905A ; CURRENT FILING DATE: 2001-04-27</p> | | | |

PRIOR APPLICATION NUMBER: US 60/200,198
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn version 3.1
SEQ ID NO 11
LENGTH: 1338
TYPE: DNA
ORGANISM: Homo sapiens
FEATURE:
NAME/KEY: misc_feature
LOCATION: (513)..(513)
OTHER INFORMATION: unsure
US-09-843-905A-11

Alignment Scores:
Pred. No.: 1,02e-168 Length: 1338
Score: 1642.00 Matches: 299
Percent Similarity: 81.4% Conservative: 42
Best Local Similarity: 71.4% Mismatches: 74
Query Match: 71.7% Indels: 4
DB: 3 Gaps: 3

US-10-041-030-4 (1-420) x US-09-843-905A-11 (1-1338)

QY 3 SerProGlyGlnGluGlnHisCysAlaProAsnLysGluProValLysTyrGlyGluLeu 22
DB TCTCCCGGT-----GAAAGATGCCGACGCCAGGAGAGCCCAATCAAGTATGTGAATC 141

QY 23 ValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgLysArgLysSerArg 42
DB 142 ATGTCCTGGGCTCAATGCTTGTCTGCAAGTGGGAGCAAGGGCCGCCGCAAGCCGC 201

QY 43 PheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHisValIle 62
DB 202 CTGGCAGCTGACGGCGCTGACAGCCCAAGGGGTGAAGCCACACGTCATGACACATC 261

QY 63 SerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerLysTyrThr 82
DB 262 TCCAGCCCGCTCTCTCCCAAGGACCTGATACCGTGTGACGACACATCTCGTATACA 321

QY 83 LeuSerArgAsnGlnThrValAlaValGluTyrThrHisAspLysAspThrAspMetPhe 102
DB 322 CTCTCCCGAGCCACTCGCTCATAGTGAGATATACATATACGACACAGACATCTTC 361

QY 103 GlnValGlyArgSerThrGluSerProLysAspPheValAlaThrAspThrLysSerGly 122
DB 382 CAGATTGCGCGCTCCACAGAGAACATGATTGACTTCGTGTAAACAGACAGTCCCGTGA 441

QY 123 SerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrLysSerArgPheAlaCysArg 142
DB 442 GGA--GGGGCTGCCAGGGCCCTTCTGCCCAGAGCACATCTCCCGCTATGCTGCCG 498

QY 143 IleValCysAspArgAsnGluProLysTyrThrAlaArgLysPheAlaAlaGlyPheAspSer 162
DB 499 ATCTCTGTACACGCGCCACCTTATCTCCCGACATCTATGCGCTGCTTCGATGCC 558

QY 163 SerLysAsnLysPheLeuGlyGlyValAlaAlaLysTyrLysAsnProAspGlyHisMet 182
DB 559 TCTAGCAACATCTTCTCTTGAAGCGAGCGGCAAAATGGGAGACCCAGATGGCGTGAAG 618

QY 183 AspGlyLeuThrThrAsnGlyValLeuValMetHisProArgLysGlyLysPheThrGlu 202
DB 619 GATGAGACTGACCAACATGAGTCTGTGATGACCCCGCAGGCGGCTTCTCCGAGAGAC 678

QY 203 SerGlnProGlyValTyrPArgGluLysSerValCysGlyAspValTyrThrLeuArgGlu 222
DB 679 TCGAGCCCGGCTGTCTGCGGAGATCTCGCTCTGTGGAAATGTATCAATTCGCGGAC 738

QY 223 ThrArgSerAlaGlnArgLysLysLeuValGlnSerGluThrAsnValLeuGlnAsp 242
DB 739 AGCCGCTCAGCCAGAGCGGGGAGCTGTGTAAGAAACAGATCCAAAGTCTGCGAGAC 798

QY 243 GlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTyrPArgThrAlaAspGlyLeuPhe 262

DB 799 GGCTCTCATGACCTGTGTGGGGCCACATGCTGTGGGACACACCGGCGGGCTGCTG 858

QY 263 HisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluLeuAsnAlaAlaArgPro 282
DB 859 CGGGCTCCACCTGAAGCACTGAGAGGCCCGAGCGGACAGAGCAAAATGACAGCGCCGCC 918

QY 283 GlnCysProValGlyLeuAsnThrLeuAlaPheProSerLysAsnArgLysGluValAla 302
DB 919 CAGTGCCTGGTGGGCTTCAAGCACTTGGGCTTCCAGGCCAGCCCGGTGGCCGACAGCG 978

QY 303 GlnGluLys---GlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHisAsn 321
DB 979 CCGGACAAACAGAGAGCCCTGGGTCTACGTCCTGGGGGACAGCTCATGCTACACAGCGC 1038

QY 322 TTPGlyHisArgSerAspThrGluAlaAsnGluArgGlnCysProMetCysArgThrVal 341
DB 1039 TGGGGCTGGCGGGGAGCGGGGCCCGGAGCGGAAATGTCTCTCTGCGCCCTGTGTG 1098

QY 342 GlyProTyrValProLeuThrLeuGlyCysGluAlaGlyPheTyrValAspAlaGlyPro 361
DB 1099 GGGCTTATGTGCTCTATGCTTGGCCAGAGGCGGCTTGTGCTGACCTTGAGCCTGGGCGG 1158

QY 362 ProThrHisAlaPheThrProCysGlyHisValCysSerGlyLysSerAlaLysTyrTrp 381
DB 1159 CTTAGCCATGCTTTGACCTTGGCCGACGCTGCTCTGAGAAAGATCCGCGCTACTGG 1218

QY 382 SerGlnIleProLeuProHisGlyTyrHisAlaPheHisAlaAlaCysProPheCysAla 401
DB 1219 GCCCAGACACCACTGGCCCAAGGACCAATGCTTTCATAGCGCGCTGCGCCCTTTGCGGG 1278

QY 402 ThrGlnLeuValGlyGluGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
DB 1279 GCCTGCTTACCGCGCAGAGATGGCTGTCGCTCATTTTCCAGGGCCCGCTGAT 1335

RESULT 6
US-10-104-047-1958
Sequence 1958, Application US/10104047
Patent No. 6943241
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. 6943241el full length cdna
FILE REFERENCE: H1-A0105
CURRENT APPLICATION NUMBER: US/10/104,047
PRIOR FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER:
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 1958
LENGTH: 2508
TYPE: DNA
ORGANISM: Homo sapiens
US-10-104-047-1958

Alignment Scores:
Pred. No.: 6.73e-151 Length: 2508
Score: 1482.50 Matches: 275
Percent Similarity: 74.2% Conservative: 36
Best Local Similarity: 65.6% Mismatches: 65
Query Match: 64.7% Indels: 43
DB: 3 Gaps: 4

US-10-041-030-4 (1-420) x US-10-104-047-1958 (1-2508)

QY 3 SerProGlyGlnGluGlnHisCysAlaProAsnLysGluProValLysTyrGlyGluLeu 22
DB 208 TCTCCCGGT-----GAAAGATGCCGACGCCAGGAGAGCCCAATCAAGTATGTGAATC 261

QY 23 ValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgLysArgLysSerArg 42
DB 261 ----- 261

QY 43 PheAlaLeuTyrIlybArgProLybAlaAsnGlyValIlybProSerThrValHisValIle 62
 Db 262 -----ATC 264
 QY 63 SerThrProGlnAlaSerIlybAlaIleSerCysIlybGlyGlnHisSerIleSerIlybThr 82
 Db 265 TCCACCCCGCTCGCTCCAGAGCACTGATACCTGGTGTACAGACACATCTGTATACA 324
 QY 83 LeuSerArgAsnGlnThrValValIlybGlyTyrThrHisAspIlybAspThrAspMetPhe 102
 Db 325 CTCTCCCGAGCACTCGTGTATAGTGGATATACATATATAGCCGACACAGACATGTTTC 384
 QY 103 GlnValGlyArgSerThrGlnSerProIleAspPheValIlybThrAspThrIleSerGly 122
 Db 385 CAGATTGGCCGCTCCACAGAACATGATGACTTCTGTATACAGACAGCTCCCGGGA 444
 QY 123 SerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArg 142
 Db 445 GGA---GGGCTGCGAGGGCCCTTTCGCCAGACACATCTCCGCTATGCTGCGCG 501
 QY 143 IleValCysAspArgAsnGlnProTyrThrAlaArgIlePheAlaAlaGlyPheAspSer 162
 Db 502 ATCTCTGTACCGCGCCACCTTATACCTGCGCATCTATGCGCTGCTTGATGAGCC 561
 QY 163 SerIlybAsnIlePheLeuGlyGlyIlybAlaAlaIlybTyrIlybAsnProAspGlyHisMet 182
 Db 562 TCTAGCAACATCTTCTTGTAGAGCCAGCGCCAAATGGCGGACCCAGATGGCTGATG 621
 QY 183 AspGlyLeuThrThrAsnGlyValIlybValIleuValMetHisProArgGlyGlyPheThrGln 202
 Db 622 GATGAGTACACACCAATGAGTCTGTGATGACCGGACGGCGGCTTCTCCAGAGAC 681
 QY 203 SerGlnProGlyValIlybArgGlnIleSerValCysGlyAspValIlybThrLeuArgGln 222
 Db 682 TCAAGCCCGGCTGTGCGGAGATCTCGTGTGAGATGTGATGACATTCGCGGAC 741
 QY 223 ThrArgSerAlaGlnGlnArgGlyIlybLeuValGlnSerGlnThrAsnValIleuGlnAsp 242
 Db 742 AGCGGCTTACAGCCAGCGCGGGGACACTGGTGGAAACAGATCCATGCTGCGAGAC 801
 QY 243 GlySerLeuIleAspLeuCysGlyValIlybThrLeuLeuTyrThrAlaAspGlyLeuPhe 262
 Db 802 GGCTCTCTATCACTGCTGTGGGGCCACACTGCTGTGGGACACCGCGGGGCTGCTG 861
 QY 263 HisThrProThrGlnIlybHisIleGlnAlaLeuArgGlnGlnIleAsnAlaAlaArgPro 282
 Db 862 CGGGCTTCCACACTGAAAGCACTGAGAGCCGACGCGACGAGGCAATGCAAGCGCGGCC 921
 QY 283 GlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgIlybGlnValIle 302
 Db 922 CAGTGGCCCGTGGGCTTCAAGCACTTGGCTTCCCGACCCGCGGCTGGCGGACGCG 981
 QY 303 GlnGlyIlyb---GlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHisAsn 321
 Db 982 CCGGACAAACAGAGCGCTGGGTCTACGTCCGCGGAGCAAGTCCATGCTACAGAGC 1041
 QY 322 TyrGlyHisArgSerAspThrGlnAlaAsnGlnArgGlnCysProMetCysArgThrVal 341
 Db 1042 TGGGGCTGCGCGGAGCGGGGCGCCGAGAGGCGCAATGCTCTGCTGCGGCTGCTG 1101
 QY 342 GlnProTyrValProLeuTyrPheGlyCysGlnAlaGlyPheTyrValAlaAspAlaGlyPro 361
 Db 1102 GGGGCTTATGCTCTATGCTTGGCCAGAGGCGCGCTTGGCTGAGACCTGGGCGG 1161
 QY 362 ProThrHisAlaPheThrProCysGlyHisValCysSerGlnIlybSerAlaIlybTyrTrp 381
 Db 1162 CCAAGCATGCTTTCAGACCTTGGCGGACAGTCTGCTGAGAGAGATGCGCGTACTG 1221
 QY 382 SerGlnIleProLeuProHisGlyTyrHisAlaPheHisAlaAlaCysProPheCysAla 401
 Db 1222 GCCCAGACACACAGCGCCACGCGACCATGCTTTCATGCGCGCTGCTGCGGCG 1281
 QY 402 ThrGlnLeuValGlyGlnGlnAsnCysIleIlybLeuIlePheGlnGlyProIleAsp 420

Db 1282 GCGTGGCTTACCGGAGCATGCGTGCATCGCGCTCATTTTCCAGGCGCGTGCAT 1338
 RESULT 7
 US-08-136-277-1
 / Sequence 1, Application US/08136277
 / Patent No. 5644045
 / GENERAL INFORMATION:
 / APPLICANT: MANDEL, Jean-Louis
 / APPLICANT: AUBOURG, Patrick
 / APPLICANT: MOSSER, Jean
 / APPLICANT: SARDE, Claude
 / TITLE OF INVENTION: X-LINKED ADRENOLEUKODYSTROPHY GENE AND
 / NUMBER OF SEQUENCES: 23
 / CORRESPONDENCE ADDRESS:
 / ADDRESSEE: Young & Thompson
 / STREET: 745 South 23rd Street
 / CITY: Arlington
 / STATE: VA
 / COUNTRY: USA
 / ZIP: 22202
 / COMPUTER READABLE FORM:
 / MEDIUM TYPE: Floppy disk
 / COMPUTER: IBM PC compatible
 / OPERATING SYSTEM: PC-DOS/MS-DOS
 / SOFTWARE: Patent Release #1.0, Version #1.25
 / CURRENT APPLICATION DATA:
 / APPLICATION NUMBER: US/08/136,277
 / FILING DATE: 15-OCT-1993
 / CLASSIFICATION: 424
 / ATTORNEY/AGENT INFORMATION:
 / NAME: PATCH, Andrew J.
 / REGISTRATION NUMBER: 32,925
 / REFERENCE/DOCKET NUMBER: B2272
 / TELECOMMUNICATION INFORMATION:
 / TELEPHONE: 703-521-2297
 / TELEFAX: 703-685-0573
 / TELEX: 248425 EMBOW
 / INFORMATION FOR SEQ ID NO: 1:
 / SEQUENCE CHARACTERISTICS:
 / LENGTH: 2750 base pairs
 / TYPE: nucleic acid
 / STRANDEDNESS: single
 / TOPOLOGY: linear
 / MOLECULE TYPE: cDNA
 / FEATURE:
 / NAME/KEY: CDS
 / LOCATION: 387..2624
 / US-08-136-277-1
 Alignment Scores:
 Pred. No.: 0.0186 Length: 2750
 Score: 120.50 Matches: 104
 Percent Similarity: 33.9% Conservative: 63
 Best Local Similarity: 21.1% Mismatches: 155
 Query Match: 5.3% Indels: 171
 DB: 2 Gaps: 26
 US-10-041-030-4 (1-420) x US-08-136-277-1 (1-2750)
 QY 2 PheSerProGlyGlnGlnGlnHisCysAlaProAsnIlybGlnProValIlybTyrGlyGln 21
 Db 1167 TTCTCGCC-----AAATTGCGGAG 1187
 QY 22 LeuValValIleuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgArgIlybSer 41
 Db 1188 CTGCTG-----GCAGAGAGCGCGCGGAGGCGG 1217
 QY 42 ArgPheAla---LeuTyrIlybArgProLybAlaAsnGlyValIlybProSerThrValHis 60
 Db 1218 GAGCTGGCTACAGACATGCGTGTGTGGCCAAC----- 1253

61 ValIleSerThrProGlnAlaSerValAlaIleSerGlyValGlnHisSerIleSer 80
 1254 -----TCGAGAGAGATCGCTTTATGTGGAGGCCAATGAGGGTGG
 81 TyrThrLeu-----SerArgAsnGlnThrValValVal 91
 1293 CTGGCCCTCGTACACGCGCTCCACACGAGACCTGGCTCGCAATGACCTCATCTTCTG 1352
 92 Gln-----TyrThrHisAspIleAspThrAspMetPheGlnValGlyAspSerThr 108
 1353 GAAGCCCTGTGGTATGTTATGCTCGAGACAGATTCCTCAAGAATATGTGTGAGCGCCCTG 1412
 109 GluSerProIleAspPheValValThrAspThrIleSerGlySerGlnAsnThrAspGlu 128
 1413 GGCCTGCTCAATGGTGGCTGTGCTCCCATCATCACTGCGACCTGCTACTAGCTACGATGCA 1472
 129 AlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgIleValCysAspArgAsn 148
 1473 GAGCCCGTGAAGAAAGACGCTTGAAAGAAAGAGAGAGAGCTGTGTAGGAGCGACCA 1532
 149 GluProTyrThr-----AlaArgIlePheAlaAlaGlyPheAspSerSerIleAsnIle 166
 1533 GAAGCCTTCACTAATGCCCCGCAACCTCTCGACAGCGCTGCAAGATGCATTGTAGCGGATC 1592
 167 -----PheLeuGlyGlnValAlaAlaValSerTyrAsnProAspGlyHisMetAsp 183
 1593 ATGTGTCGTGCACA--GGAAGTGAACGAGACTGAGCTGCTTACACAGCCCGGTGCACGAGAT 1651
 184 -----GlyLeu-----185
 1652 GTTCCAGGATATTGAAAGATGTTACAGCGCTGTCATCTTCAAGAGCCCGGACGAGCTAGGA 1711
 186 ThrThrAsnGlyValLeu-----ValMetHisProArgGlyGlyPheThrGlu 201
 1712 CGCTCAGCGGGGTCTGGACCATAGCGCGGTCTGTGTCCTGGTGGAGGGCCCCCTGAA 1771
 202 GluSerGlnProGlyValTyrArgGlnIleSerValCysGlyAspValTyrThrLeuArg 221
 1772 GATCCGAGGCCAGGT-----GATGATGTGAAACAGGAGATCATCTGCGA 1816
 222 Glu--ThrArgSerAlaGlnGlnArgGlyValLeuValGluSerGlnThrAsnVal--Le 240
 1817 GAACATCCCCATCGTCACAGCCCTCAGAGAGGTGTGTGTCGACAGCTCAACATCAGGCT 1876
 240 uGlnAspGly-----SerLeuIle 246
 1877 GAGGAAAGACGATGCATGCTCATACAGAGCCCAATGCTCGCGCAAGAGACTCCCTGT 1936
 246 AspLeuCysGlyAlaThrLeuLeuTyrArgThrAlaAspGlyLeuPheHisThrProTh 266
 1937 CCGAGATCCGTGGGTGG-----CTCTGCGCCACGTAACGTAGTGTCTCTCAAGACCCC 1990
 266 rGlnLysHisIleGlnAlaLeuAspGlnGlnIleAsnAlaAlaArgProGlnCysProV 286
 1991 ACCCCAGCGCATGTTCTACATCCGCA-----AGGCTTACATGTCTGT 2035
 286 IGI-----LeuAsnThrLeuAlaPheProSerIleAsnArgGlyGlnValGluI 304
 2036 GGGCTCCCTGGGTATGACAGAGTATACCG-----GACTCAGTGGAGGA 2080
 304 uLysGlnProTyrAlaTyr-----LeuSerCysGlyHisIleVal 316
 2081 CATGCAAGAGAAAGGCTACTCGAGACAGACACTGGAGCCATTCCTGAGCTGTGACCT 2140
 316 IHis-----GlyTyrHis-----AsnTyrGlnHis--Ar 325
 2141 GACACCAATCTCTGACGCGGAGAGAGAGGTGGAGGCTATGTGTATCTGAAAGACGTCT 2200
 325 gSerAspThrGlu--AlaAsnGlnArgGlnCysProMetCysArgThrValGlyProTyrV 345
 2201 GTCCGGGTGGCAGAAAGACAGAAATTCGGCATGGCCCGCATGTTCTTACACAGGCCCAAGTA 2266
 345 aProLeuTyrLeuGlnCysGlnAlaGlyPheTyrValAspAlaGlyPro----- 361

| | | | | | | | | |
|----|------|--|-------|------------|-------|-------|--------------|------|
| Db | 2261 | | ::: | | | | CGGAGACATCGA | 2299 |
| Qy | 362 | ----- | ----- | ----- | ----- | ----- | ProThr | 364 |
| Db | 2300 | CGTGAAGGCAAGATCTTTCAGAGCGGCCAGAGACGGGCAATTGCCTCTCTCAATAC | | | | ::: | | 2359 |
| Qy | 364 | IsaIaPheThrProCys61y | ----- | -HisValCys | | ----- | -Serg | 375 |
| Db | 2360 | CCACCGGCGCCCTCCTGTGGAAATACACACACATTTGCTACATTCATGGGGAGGGCGG | | | | ::: | | 2419 |
| Qy | 375 | IuIySeraIaIySyrTrpSergIuIePro | | | | | | 385 |
| Db | 2420 | CTGGAAGTTCAGAAAGCTGAGCTGACCTGACGCGCC | | ::: | ::: | ::: | | 2451 |

```

US-08-479-403-1
/ Sequence 1, Application US/08479403
/ Patent No. 586303
/ GENERAL INFORMATION:
/ APPLICANT: MANDEL, Jean-Louis
/ APPLICANT: AUBOURG, Patrick
/ APPLICANT: MOSSER, Jean
/ APPLICANT: SAUDE, Claude
/ TITLE OF INVENTION: X-LINKED ADRENOLEUKODYSTROPHY GE
/ TITLE OF INVENTION: CORRESPONDING PROTEIN
/ NUMBER OF SEQUENCES: 23
/ CORRESPONDENCE ADDRESSES:
/ ADDRESSEE: Young & Thompson
/ STREET: 745 South 23rd Street
/ CITY: Arlington
/ STATE: VA
/ COUNTRY: USA
/ ZIP: 22202
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patentin Release #1.0, Version #1.25
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/08/479,403
/ FILING DATE: 07-JUN-1995
/ CLASSIFICATION: 514
/ ATTORNEY/AGENT INFORMATION:
/ NAME: PATCH, Andrew J.
/ REGISTRATION NUMBER: 32,925
/ REFERENCE/DOCKET NUMBER: B2722DIV
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 703-521-2297
/ TELEFAX: 703-685-0573
/ TELEX: 248425 EMBON
/ INFORMATION FOR SEQ ID NO: 1:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2750 base pairs
/ TYPE: nucleic acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: cDNA
/ FEATURE:
/ NAME/KEY: CDS
/ LOCATION: 387..2624
/ US-08-479-403-1

Alignment Scores:
Pred. No.: 0.0186 Length: 2750
Score: 120.50 Matches: 104
Percent Similarity: 33.9% Conservative: 63
Best Local Similarity: 21.1% Mismatches: 155
Query Match: 5.3% Indels: 171
DB: 2 Gaps: 26
US-10-041-030-4 (1-420) x US-08-479-403-1 (1-2750)

```

```

Qy 2 PhaserProgluYngluYnHsCyaaLaPbRoAnlysgluProValuYtyrYglu 21
Db 1167 TTCTCGCC-----AGTTCTGGGAG 1187
Qy 22 LeuValValLeuGlyYraEnglyLaLeuProAnlyAspArgGlyAraGlySer 41
Db 1188 CTGGTG-----GCAGAGGCGCGCGGAGGG 1217
Qy 42 ArgPheAla---LeuYtyrLysArgProLyAlaAnlyValuYsProSerThrValHis 60
Db 1218 GAGCTGGCTACATGACCTGCGTGTGGCCAAC----- 1253
Qy 61 ValIleSerThrProGlnAlaSerLyAlaIleSerCylysglyGlnHisSerIleSer 80
Db 1254 -----TCGAGAGATGCGCTTCTATGGGGGCGCATGAGTGAG 1292
Qy 81 TyrThrLeu-----SerArgAnlyThrValValu 91
Db 1293 CTGGCCCTGTACAGCGCTCTACAGAGACCTGGCTCGCATCACTCATCTTCTG 1352
Qy 92 Glu-----TyrThrHisAspLyAspThrAspMetPheGlnValGlyArgSerThr 108
Db 1353 GAACGCTGTGTGTATGTATGTCTGAGACAGTCTCTCAAGATATGTGTGAGCGCTCG 1412
Qy 109 GluSerProIleAspPheValValThrAspThrIleSerGlySerGlnAsnThrAspGlu 128
Db 1413 GCGCTGTATGTGTGTCTCCCATCATCATCTGCTCTACAGAGTCAAGTCA 1472
Qy 129 AlAGlnIleThrGlnSerThrIleSerArgPheAlaCyaaGlnIleValCyaaAspAsn 148
Db 1473 GAGCGCTGTAAAGAGCGCTTGAAAGAGAGAGAGAGAGCTGTGAGCGAGCGCA 1532
Qy 149 GluProtyrThr-----AlaArgIlePheAlaIleGlyPheAspSerSerLyAsnIle 166
Db 1533 GAAGCTTCACTATTGCGCGCACTCTGACAGCGCTGAGTGCATTTGAGCGGATC 1592
Qy 167 -----PheLuglyGlyValAlaIleTyryrLysAsnProAspGlyHisMetAsp 183
Db 1593 ATGTCTGTCAAA--GGAAGTGAAGAGCTGTGCTGCTACAGACCGGCTGACAGAT 1651
Qy 184 -----GlyLeu----- 185
Db 1652 GTTCCAGATATTCAGATGTTCAAGCGCTGCACTTCAAGAGCGCCAGAGCTAAGAG 1711
Qy 186 ThrThrAsnGlyValLeu-----ValMetHisProAspArgGlyPheThrGlu 201
Db 1712 CCGTCAAGCGCGGTGAGACCATAGCGCGGTGTGTGTGTGAGAGCGCCCTGAA 1771
Qy 202 GluSerGlnProGlyValTTPArgGluIleSerValCyaaGlyAspValTyrThrLeuArg 221
Db 1772 GATCCAGGCGCAGGT-----GTTGATGTGAGACAGGAGATCATCTCGA 1816
Qy 222 Glu-ThrArgSerAlaGlnIleArgGlyLyLeuValGluSerGlnThrAsnVal---Le 240
Db 1817 GAACATCCCATGTGTACAGCCCTCAGAGAGAGTGTGTGTGCGCCAGCTCAACATCAGGT 1876
Qy 240 uGlnAspGly-----SerLeu 246
Db 1877 GAGAGAGAGGATGATCTGTCTCATACAGGCCCATGCGTGGCGGAGAGCTCCCTGT 1936
Qy 246 eaAspLeuCyaaGlyAlaThrLeuLeuTTPArgThrAlaAspGlyLeuPheHisThrProth 266
Db 1937 CCGGATCTGTGGTGG-----CTCTGACCATCAAGTGTGTGTCTACAAAGCCCTC 1990
Qy 266 rGlnLyHisIleGlnAlaLeuArgGlnGluIleAsnAlaAlaArgProGlnCyProva 286
Db 1991 ACCCGAGCGATGTTCAATCCGAG-----AGGCTTCAATGTCTGT 2035
Qy 286 lGly-----LeuAsnThrLeuAlaPheProSerIleAsnArgLyGluValGlu 304
Db 2036 GGGCTCTCTGCTGACGAGGATCAACCG-----GACTCAGTGAGAGA 2080
Qy 304 uLyGlnProTTPAlaYr-----LeuSerCyaaGlyHisVa 316

```

```

Db 2081 CATGCAAGAGAGGCTACTCGAGAGGACCTGAAAGCATCTGACCTGCACT 2140
Qy 316 lHis-----GlyTyHis-----AsnTyrGlyHis---Ar 325
Db 2141 GCACCAATCTTCACCGGAGGAGGAGGTGGAGGCTATGTGTGATGGAGAGACCTCT 2200
Qy 325 gSerAspThrGlu-AlaAsnGluArgGlyCyaaProMetCyaaThrValGlyProTyrv 345
Db 2201 GTGGGTGGGGAAGACAGAGATTCGCATGGCCCATGTTCTTACACAGGCGCAAGTA 2260
Qy 345 aLProLeuTTPLeuGlyCyaaGluAlaGlyPheTyValAspAlaGlyPro----- 361
Db 2261 CGCCCTCTGATGAATGACACCGCG-----CGTGAGCATCGA 2299
Qy 362 -----ProThr 364
Db 2300 CGTGAAGCGAGATCTTCCAGCGGCGCAAGAGCGGCGCATTCCTGCTTCCATCAC 2359
Qy 364 lAlaPheThrProCyaaGly-----HisValCyaa-----SerG 375
Db 2360 CCAACGCGCTCTCTGTGAAATACACACACACTTGTCAAGTTCAGTGGAGGCGG 2419
Qy 375 lLysSerAlaLyTyryrTTPSerGlnIlePro 385
Db 2420 CTGAAGTTCGAGAGACTGAGCTGACGCTGCC 2451

RESULT 9
US-08-835-734-1
Sequence 1, Application US/08835734
Patent No. 6013769
GENERAL INFORMATION:
APPLICANT: MANDEL, Jean-Louis
APPLICANT: AUBOURG, Patrick
APPLICANT: MOSSER, Jean
APPLICANT: SARDE, Claude
TITLE OF INVENTION: X-LINKED ADRENOLEUKODYSTROPHY GENE AND
TITLE OF INVENTION: CORRESPONDING PROTEIN
NUMBER OF SEQUENCES: 23
CORRESPONDENCE ADDRESS:
ADDRESSER: Young & Thompson
STREET: 745 South 23rd Street
CITY: Arlington
STATE: VA
COUNTRY: USA
ZIP: 22202
COMPUTER READABLE FORM:
MEDIUM TYPE: Floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: PatentIn Release #1.0, Version #1.25
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/08/835,734
FILING DATE:
CLASSIFICATION: 536
PRIOR APPLICATION DATA:
APPLICATION NUMBER: US 08/479,403
FILING DATE: 07-JUN-1995
ATTORNEY/AGENT INFORMATION:
NAME: PATCH, Andrew J.
REGISTRATION NUMBER: 32,925
REFERENCE/DOCKET NUMBER: B272721V
TELECOMMUNICATION INFORMATION:
TELEPHONE: 703-521-2297
TELEFAX: 703-685-0573
TELEX: 248425 EMBON
INFORMATION FOR SBO ID NO: 1:
SEQUENCE CHARACTERISTICS:
LENGTH: 2750 base pairs
TYPE: nucleic acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: cDNA

```


Query Match: 5.0% Indels: 129
DB: 3 Gaps: 21
US-10-041-030-4 (1-420) x US-09-103-840A-2 (1-4403765)
QY 4 ProGlyGlnGlu-----GluHieCyAlaProAnlyGluProValLysTyr 19
Db 1250835 CCGTGGCAACAAAGGCCAGCATGACATGATCGCGGACACCATGTGCC----- 1250785
QY 20 GYGluLeuValValLeuGlyTyrAanglyAlaLeuProAnnglyAaPaArgGlyArgArg 39
Db 1250784 GGTGGCGCGCTGCAATTCGATTACAC-----CGAGGTGGCGGT 1250746
QY 40 LysSerArgPheAlaLeuTyrLysArgProLysAlaAnnglyValLysPro----- 56
Db 1250745 -----GGCTATCTCGCGCGCGCGCATCCAGATCCCGCGCATCGGATGA 1250698
QY 57 -----SerThyValHieValLieserThPro 65
Db 1250697 CGTGTGATCGATGAGTGTGCTGGAACAGAAACCTCGTGACGTTGCGCGGTGCC 1250638
QY 66 GluAlaSerLysAlaLieserCyLysGlyGlnHieSerIle-----Ser 80
Db 1250637 GAATCCGCTGTGTGTCAGCGGTGCGGTGAGGACCCAGTCGATGCGCGCATGATGA 1250578
QY 81 TyrThreLysSerArg-AangInThValValGlu----- 92
Db 1250577 CAGCATTCGAGACGTGGGCGATGACCGCGGTGAAACGATCTTGCAGCGCGGTGC 1250518
QY 93 -----TyrThriHisAaP-LysAaPThrAaPMeT-----PheGlnValGlyArgSerThx 109
Db 1250517 CGCGCATTCGCGCCACACACACACGCGCATTCACACGATGATGCGGCGCATGCTGCG 1250458
QY 109 LysEProLleAaPPhaValValThraPThrLieserGlySerGlnAaThraPelu- 128
Db 1250457 CAGGCGGTGCCATTCACACGACCGGTATGCCAGTGTGGCGCGCATCGAGCGCGT 1250398
QY 129 -----AlaGlnLleThGlnSerThr- 135
Db 1250397 ATTGCGCTCCTGGAATCAGTCCGCGCGGTGCTTATCGCACCATCGGCTCGACGA 1250338
QY 136 -----LieserArgPheAlaCyAaGlyLeValCyAaPaArgAangInProTyrThx 153
Db 1250337 CCAAGCGGTACCGCGGTGTGTGTGCAAGCGATGTATGCGCAACCTCACCG----- 1250285
QY 153 laaArgLlePheAlaAlaGlyPheAaPserSerLysAaLlePheLeuGly-----GluL 171
Db 1250284 --CGAACTCG-GGCGCGGTGTGTGTCTTCACCAACCGCATCACCGAGCCAAACGAAC 1250228
QY 171 ysaAlaAlaLysTyrLysAaPProAaPArgLysMeTAsaPArgLysLeuThThraAnnglyValL 191
Db 1250227 CGTTGCGGAATGGCTA-----CGCGCGCGCGCAAGCGAGACGTG-----GTGTCCGGAATGG 1250174
QY 191 euValMeThAaPProAaArgGlyGlyPheThrgInGluSerGlnProGlyValTyrArgGlu 211
Db 1250173 TCGCGGTGCGACCGATACCGCGCTGGCGACAG-----CAACCGCGGTCTTACAGAAC 1250117
QY 211 leSerValCyGlyAaPValTyrThreAaArgGluThraArgSerAlaGlnInArgGlyL 231
Db 1250116 TG-----ATGCGCGCGCGCGCGCGGTGAGCGCATGCGACG 1250081
QY 231 ysaLeuValGluSerGluThraAnValLeuGlnAaPArgLys-----LeuLleAaPLeu 249
Db 1250080 GCGACGTTCAGAGATGCAATTCACCGTGGAGAGACAGCCAGCTGTGCTGACAGACC 1250021
QY 249 ysaGlyAlaThreLysLeuTyrArgThraAlaAaArgGlyLeuPheHieThProThrgInLysH 269
Db 1250020 GCGGGGCGGAA-----CGCTCGCGCGAGCGCGCGCGGTGCACTGCGCGGTGCACTGC 1249970
QY 269 lsaLleGlnAlaLeuArgGlnLysLleAaAlaAlaArgProGlnCyAaProValGlyLeuA 289
Db 1249969 ATCAGAGGAGCTATGACGACACCGAGACACTGCGCGGTGATCTCGACGACGACATCG 1249910

QY 289 snThreAlaPheProSerIle-AaPaArgGluValValGluGluLysGlnProTyr 308
Db 1249909 AGACTTGTGCTACGGCGGTGCGTGCAGACCGAAA----- 1249877
QY 309 AlaTyrLeuSerCyGlyHieValHieGlyTyrHieAaPTrpGlyHieAaPserAaPThr 328
Db 1249876 -----CAGGTTGCTGTGACCGC----- 1249859
QY 329 GluAlaAaAngLysArgGluCySPrometCyAaGThValGlyProTyrValProLeuTyr 348
Db 1249858 ---TTTGGCCAAAGGCGCTGCGCGTTCGCGCGCGGTGATTCG----- 1249817
QY 349 LeuGlyCyGlyAlaGlyPheTyrValAaPaArgLysProProThriAlaAaPThrPro 368
Db 1249816 -----GGAACCGCTTACACCGAGTGCAGACG 1249790
QY 369 CyGlyHieValCySerGluLysSerAlaLysTyrTyrSerGlnLle 384
Db 1249789 CGCTGACGCTGCGGACCGGGCGACCGGTCACTTGTGTGCGAGATC 1249742
RESULT 11
US-09-103-840A-1/C
; Sequence 1, Application US/09103840A
; Patent No. 6294328
; GENERAL INFORMATION:
; APPLICANT: FLEISCHMAN, Robert D.
; APPLICANT: WHITE, Owen R.
; APPLICANT: FRASER, Claire M.
; APPLICANT: VENTER, John C.
; TITLE OF INVENTION: DNA SEQUENCES FOR STRAIN ANALYSIS IN MYCOBACTERIUM
; FILE REFERENCE: 24366-20007.00
; CURRENT APPLICATION NUMBER: US/09/103,840A
; NUMBER OF SEQ ID NOS: 2
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 1
; LENGTH: 4411529
; TYPE: DNA
; ORGANISM: Mycobacterium tuberculosis
; OTHER INFORMATION: H37Rv
US-09-103-840A-1
Alignment Scores:
Pred. No.: 1,7e+04 Length: 4411529
Score: 113.50 Matches: 98
Percent Similarity: 34.0% Conservative: 51
Best Local Similarity: 22.4% Mismatches: 161
Query Match: 5.0% Indels: 129
Gaps: 21
US-10-041-030-4 (1-420) x US-09-103-840A-1 (1-4411529)
QY 4 ProGlyGlnGlu-----GluHieCyAlaProAnlyGluProValLysTyr 19
Db 1251307 CCGGGCAACAAAGGCCAGCATGACATGATCGCGGACACCATGTGCC----- 1251257
QY 20 GYGluLeuValValLeuGlyTyrAanglyAlaLeuProAnnglyAaPaArgGlyArgArg 39
Db 1251256 GGTGGCGCGCTGCAATTCGATTACAC-----CGAGGTGGCGGT 1251218
QY 40 LysSerArgPheAlaLeuTyrLysArgProLysAlaAnnglyValLysPro----- 56
Db 1251217 -----GGCTATCTCGCGCGCGCGCATCCAGATCCCGCATTCGGATGA 1251170
QY 57 -----SerThyValHieValLieserThPro 65
Db 1251169 CGTGTGATCGATGAGTGTGCTGGAACAGAAACCTCGTGACGTTGCGCGGTGCC 1251110
QY 66 GluAlaSerLysAlaLieserCyLysGlyGlnHieSerIle-----Ser 80
Db 1251109 GAATCCGCTGTGTGTCAGCGGTGCGGTGAGGACCCAGTCGATGCGCGCATGATGA 1251050

```

Qy      81 TyrThrLeuSerArg-AsnGlnThrValValGlu-----92
Db      1251049 CAGAGTTCTGGAGTGGGCGATGACCGAGCGGCTGACGAGTACTTCCAGCGCGGTGC 1250990
Qy      93 -----TyrThrHisAsp-LysAspThrAspMet---PheGlnValGlyArgSerThrc 109
Db      1250989 CGCGGACTTGGCCCGACGACGACGACGCGGATTCACACGACGATGATGCGGCGATCGCG 1250930
Qy      109 LuserProIleAspPheValValThrAspThrIleSerGlySerGlnAsnThrAspGlu- 128
Db      1250929 CAGCGCGGTCCTCATTAACGACGACCGGATGCGGCGGCGGCGGCGGCGGCGGCGG 1250870
Qy      129 -----AlaGlnIleThrGlnSerThr- 135
Db      1250869 ATTGCGCTCTGGAGTACCTGCGCGCGGCTGCTATGCGGACCATGACGCGCTTCGACGA 1250810
Qy      136 -----IleSerArgPheAlaCysArgIleValCysAspArgAsnGluProTyrThra 153
Db      1250809 CCAAGCGGTAACGCGCGGTGGTGGAGGCGGATGATGCGGACCTCAACCG----- 1250757
Qy      153 IaaArgIlePheAlaAlaGlyPheAspSerSerIysAsnIlePheLeuGly-----GluL 171
Db      1250756 --CGAAGCTCG--GGCGCGCGGCGTCTGTCTCAGCGCAACCGGATCAGCGGACCAAC 1250700
Qy      171 YaaIaAlaIuYrTrpLysAsnProAspGlyHisMetAspGlyLeuThrThrAsnGlyValL 191
Db      1250699 CGTTCCGCGGATGCTA---CCGCGCGCGCAAGGCGGACGACGAGT---GTGTCGCGATTGG 1250646
Qy      191 euValMetHisProArgGlyGlyPheThrGlnGlnSerGlnProGlyValIleTyrArgGlu 211
Db      1250645 TCGCGCGCGACCGATACCGCGCGCGCGGACGACG---CAGCGCGCGGTCTACGACCAAC 1250589
Qy      211 IeSerValCysGlyAspValTyrThrLeuArgGlnThrArgSerAlaGlnIleArgGlyL 231
Db      1250588 TG-----ATGGCGCGCGCGCGCGGATGTTGCGGCGATGCGCAG 1250553
Qy      231 YsLeuValGlnSerGlnThrAsnValLeuGlnAspGlySer-----LeuIleAspLeuc 249
Db      1250552 GCGACGTCGACGAGATGAAATTCACCGTGAAGACGACGACGCGTGTGCTGTGACGACC 1250493
Qy      249 YsGlyAlaThrLeuThrArgThrAlaAspGlyLeuPheHisThrProThrGlnLysH 269
Db      1250492 GCGCGCGCGGAA-----CGCTCGCGCGGACGCGCGCGCGCGCGCGCTGCACTGC 1250442
Qy      269 IsIleGlnAlaLeuArgGlnGlnIleAsnAlaAlaArgProGlnCysProValGlyLeuA 289
Db      1250441 ATCAGGAGGAGCTCATGACGACGACGACGACGCGCGGCGTGTGCTGCGACGACATCG 1250382
Qy      289 snThrLeuAlaPheProSerIle-AsnArgLysGluValValGluGlyLysGlnProTyr 308
Db      1250381 AGACTTGTCTACGCGCGCTGCTGCGACGCGGAA----- 1250349
Qy      309 AlaTyrIleuSerCysGlyHisValHisGlyTyrHisAsnTyrGlnHisArgSerAspThr 328
Db      1250348 -----CAGCGTGTGCTGCGACGCG----- 1250331
Qy      329 GlnIlaAsnGlnArgGluCysProMetCysArgThrValGlyProTyrValProLeuTyr 348
Db      1250330 ---TTTGGCGCAAGGCGCTGCGCGCTGCGCGCGCGGCGTATCG----- 1250289
Qy      349 LeuGlyCysGluAlaGlyPheTyrValAspAlaGlyProProThrHisAlaPheThrPro 368
Db      1250288 -----CGACCGCTTACACGAGGCGGTCGACGAG 1250262
Qy      369 CysGlyHisValCysSerGluLysSerAlaLysIleTyrProSerGlnIle 384
Db      1250261 CGTGAAGCGTGGCGGACCGGCGGCGGCGGATCTTGTGTCGAGATC 1250214

```

```

; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14397
; LENGTH: 1590
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
; US-09-252-991A-14397

Alignment Scores:
Pred. No.: 0.137 Length: 1590
Score: 109.00 Matches: 107
Percent Similarity: 30.1% Conservative: 47
Best Local Similarity: 20.9% Mismatches: 177
Query Match: 4.8% Indels: 182
DB: 3 Gaps: 25

US-10-041-030-4 (1-420) x US-09-252-991A-14397 (1-1590)
Qy      24 ValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArg-----GlyArgArgLys 40
Db      1477 CTGCTGGCGCGGATGCGGCTGCTCA---GGGAAAGCGGACGCTGCGGCTGCGGAAACG 1421
Qy      41 SerArgPheAlaLeuTyrTrpArg-----ProLysAlaAsnGlyValLysProSer 57
Db      1420 CTTCGTCGCGCGCTGCATATCGTTTGGAAATTCACCTGCACATGAGTGGCTGACCGA 1361
Qy      58 ThrValHisValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHis 77
Db      1360 ACTCGCGCGCGCTTCCGCGATCTCCCGTCCCTGTCGCGCGCGCGCTGATCAGCGGAT 1301
Qy      78 -----SerIleSerTyrThrLeuSerArg----- 85
Db      1300 CCAAGTACCTGCTGCGCGCGGACGATCCCTGCGCGGACGCTATGCGTGTCCGACAC 1241
Qy      86 -----Asn 86
Db      1240 CTGCGCGCGCGCACTGCTGACCAACCGCGCGGACGAGCGCGCGCGGAAAC 1181
Qy      87 GlnThrValValValGluTyrThrHisAspLysAspThrAspMetPheGlnValGlyArg 106
Db      1180 CGATTCGCGCGCGTACCGGATCGCCGACGAGCGGAAACGCTGCGCGCTGTGACTCGG 1121
Qy      107 SerThrGlnSerProIleAspPheValValThrAspThrIleSerGlySerGln-AsnTh 126
Db      1120 ACCGC-AGCGGTCCGATGATTCGTATC-----GAATACGAAAAAAGGAAACCA 1068
Qy      126 IAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAlaCysArgIleValCysAs 146
Db      1067 TAGCAAGACGTCGACACTGAGGCGGATTCCTCCCGC----- 1031
Qy      146 PArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPheAspSerSerLysAsnI 166
Db      1030 -----GGTCAACCACTTCAACGCG 1008
Qy      166 ePheLeuGlyGlyAlaAlaLysTrpLysAsnProAspGlyHisMetAspGlyLeuTh 186
Db      1007 ATTTCCTGTAACCTGAGGA-----AACCAACGCTCATCTC----- 968
Qy      186 TrpAsnGlyValLeuValMetHisProArgGlyGlyPheThrGlnGlnSerGlnProG 206
Db      967 -----CAACCTCG 960
Qy      206 YValTrpArgGlnIleSerValCysGlyAspValTyrThrLeuArgGlnThrArgSerAl 226

```

RESULT 12
 US-09-252-991A-14397/C
 Sequence 14397, Application US/09252991A
 Patent No. 6551795
 GENERAL INFORMATION:

```

Db      959 TGGCGACGCGGTCTCCG---GCTGGTGTATGCGGACGCTGCGGAGAACCTTCGC 903
Qy      226 aGInGInArGgLy-LyLeuValGluSerGluThrAnValIleuGlnAsp-----G 243
Db      902 TGACCATCGAGAGAAATG-----GCCATCAGAGAGTCGCCAAGATGCTCCGGCG 849
Qy      243 lYserIleuIleApleuCyseGly-----A 251
Db      848 GCCGCGTGGCGGTATCTGCGGGATTCGCGAGTTCAACAGCGCGGCGCGAGAGGTGG 789
Qy      251 laThrIleuLeuTPArGThr---AlaArGlyLeuPheHleThrPro----- 265
Db      788 CCAAGCGGTGAGCGGCGGTGCGCGCTCGACGCAATGATGATCCGCGCTGTCTATT 729
Qy      266 -----ThrGlnLysHleIleGluAlaLeuArGInGluIle----- 277
Db      728 CCTTCAGCGCGACGAGACCGCGGACATTTCCGACGCGTGGCGAGCGGACCGACTGC 669
Qy      278 -----AenAlaAlaArGProGlnCyArPro 286
Db      668 CGGTGATGCTACACACACCGCGCATCTACCGGAGACGTCACCCCGGACATCTCG 609
Qy      286 aIGlyLeu-----AenThrIleuAlaPhe-----ProSerIleAenArGly 300
Db      608 TTTCCCTGGCGCGACCTGAGAACATCGTCTGCTTCAAGACAGCTCCGCGACACCGCC 549
Qy      300 luValValIGluGluLysGlnPro-TPrAlaTyrlleuSerCyseGlyHleValHleGlyTyr 319
Db      548 GCTTCATCGACGTCACACACAGGTGGCGGACGCTTCATCTCTTCG---CGGAGCTCG 492
Qy      320 HlaAenTPrGlyHleArGSer-----AsPrThrIuAlaAenGluArGlyCyArPro 336
Db      491 ACAGCGGTGCTCGAAGAGCGTGGGAGGTGGTCCAGAGGTGATCTCGGAAATGTCCA 432
Qy      337 MetCyS----- 338
Db      431 ACAGTTCGCGAAGAGAGGAGACATCTTCGCGTGGCGAGCGCGGCTTCGCG 372
Qy      339 -----ArGThrValIGlyProTyrlValProleu 347
Db      371 AGCGGATCGCGATTCAGATGCTGATGCCGATCTTCATCTCGACCGCGCTCCGACC 312
Qy      348 TrPLeuGlyCyseGluAlaGlyPheTyrlValAraAlaGlyPro----- 361
Db      311 TGG-----TGACGTGATCATAGCTCGCGAGAACTCGCGGTGGGCGAGCGCTCA 256
Qy      362 ProThrIleAlaPheThrProCyseGlyHleValCyseSerGluLysSerAlaLysTyrlTrp 381
Db      257 CCGGTCAACGCGCGCTGGCGCTGCGCGCGCGCGATCGCGACAGCTCGAGCGGATCATGG 198
Qy      382 SerGlnIleProleuProHleGlyThrHleAlaPheHleAlaLys-----Pro 398
Db      197 CCAAGG---CCGTGGCGAGCGCGCGCGCTCGCGAGCATCGCTGTAGCCGACGCGCC 141
Qy      399 PheCyseAlaThrGlnLeuValIGlyGln 408
Db      140 GCCCTGCGCGCTCGCTCCCGGAGAGCGG 111

RESULT 13
US-09-252-991A-14280
; Sequence 14280, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27

```

```

; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 14280
; LENGTH: 3801
; TYPE: DNA
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-14280

Alignment Scores:
Pred. No.: 0.562 Length: 3801
Score: 109.00 Matches: 107
Percent Similarity: 30.1% Conservative: 47
Best Local Similarity: 20.9% Mismatches: 177
Query Match: 4.8% Indels: 182
DB: 3 Gaps: 25

US-10-041-030-4 (1-420) x US-09-252-991A-14280 (1-3801)
Qy      24 ValIleuGlyTyrlAenGlyAlaLeuProAenGlyAraPrg-----GlyArGArGly 40
Db      802 CTGCTCGGCGCGGATGCGGCTCTCA---GGCGAAAGGCGCACTGCGGGTGGCGAAG 858
Qy      41 SerArGpHeAlaLeuTyrlAraPrg-----ProLysAlaAenGlyValLysProSer 57
Db      859 CTTCGTGGGCGCTCTCATATCGTGTTCGAAATTCACCTGCCATCATGAGTCTGACCGA 918
Qy      58 ThrValHleValIleSerThrProGlnAlaSerLysAlaIleSerCyseGlyGlnHle 77
Db      919 ACTGCGCGCGCTTCGCGCATCTCCCGTCCCTGCTCGGCGGCGCGCTGATCACCGGANT 978
Qy      78 -----SerIleSerTyrlThrIleuSerArG----- 85
Db      979 CCACAGTACTGCTGCGCGCGGACGATCTCGCGCAAGGCTATCGCTGCGACAC 1038
Qy      86 -----Aen 86
Db      1039 CTGCGCGCGGCGCTGACACCGCGCGGACGAGCGAGCGCGCGGAGAAC 1098
Qy      87 GlnThrValValValGluTyrlThrHleAraPlyAraPhePheGlnValGlyArG 106
Db      1099 CGATCCGCGCGCGATCGCGATCGCCACGACGCGGAAACGCTCGCGCTGACTCGG 1158
Qy      107 SerThrIleuSerProIleAraPheValIleThrAraPheThrIleSerGlySerGln-AenTh 126
Db      1159 ACCGC-AGCGGTCCGATGATTCGTATC-----GAATATCAAGAAAGAGAACCA 1211
Qy      126 raPrgIuAlaGlnIleThrGlnSerThrIleSerArGpHeAlaCyseArGlyValCyse 146
Db      1212 TGAAGCAAGCACGTCACTGAGAGCGGCTATTCGCGC----- 1248
Qy      146 PArGAnGluProTyrlThrAlaGlyIlePheAlaIleGlyPheAraPheSerLysAenI 166
Db      1249 -----GATCACACCAATTCACAGCGG 1271
Qy      166 PheLeuGlyGluLysAlaAlaLysTyrlLysAraProAraPlyHleMetAraPlyLeuTh 186
Db      1272 ATTTCTCGTGAACCTGAGGA-----AACCCACAGCTCATCTC----- 1311
Qy      186 rThraenGlyValLeuValMetHleProArGlyGlyPheThrGluGluSerGlnProG 206
Db      1312 -----CAACCTGG 1319
Qy      206 yAlTrPaArGluIleSerValCyseGlyAraPlyTyrlThrIleuArGluThrArGserAl 226
Db      1320 TGGCGACGCGGTCTCCG---GCTGATGATTCGCGAGCGCTGCGAGAACACTTCGC 1376
Qy      226 aGInGInArGgLy-LyLeuValGluSerGluThrAnValIleuGlnAsp-----G 243
Db      1377 TGACCATCGAGAGAAATG-----GCCGTACGAGGTTCGCAAGATGCTCCGGCG 1430
Qy      243 lYserIleuIleApleuCyseGly-----A 251
Db      1431 GCGCGTGGCGGTATCTGCGGATCGCGAGTTCAACAGCGCGGCGCGAGAGGTGG 1490

```



```

QY 251 lathrleuLeuTprgthr---AlaaspGlyLeuPheH1sthrPro----- 265
DB 1491 CCAAGCGGTGGAGCGGCTGGCGTGCAGCGCATCTGATGAGCGCGCTGCTATT 1550
QY 266 -----ThrGlnYehH1leGluAlaLeuArgGlnGlu1----- 277
DB 1551 CCTCAAGCGCGCAGAGACCGCGAGCATTTCCGACGCGTGGCGGACCGGACCGACTGC 1610
QY 278 -----AsnAlaAlaArgProGlnCysArgPro 286
DB 1611 CGGTGATGTGTACAACAACCGCGCGATCTACCGGAAAGCATCCACCGGACATCGTG 1670
QY 286 alGlyLeu-----AsnThrLeuAlaPhe-----ProSerTleAsnArgGlyes 300
DB 1671 TTTCCTGGCGCGACCTGCGAGAAATCTGTCTTCAAGACAGCTCGCGGACACCGCGCC 1730
QY 300 luvAlaValGlnGluYsglnPro-TPrAlaTyrLeuSerCysGlnYlnH1sglYr 319
DB 1731 GCTTCATCGACGTAACGAGAACAGGTGGCGAGCGCTTCATCTCTCG-----CGGCTCG 1787
QY 320 H1AsnTPrGlnH1sArgSer-----AspThrGluAlaAsnGluArgGluCysPro 336
DB 1788 ACAGACGTGTCTCTGAAAGCGTGCGGTGCTCCAGGCGTGAATCTCGGAAATGTCCA 1847
QY 337 MetCys----- 338
DB 1848 AGCTGTCTCCGAAGAAAGCGGACATCTTCGCGCTGGCCAGGCGCGCGCTTCGCGG 1907
QY 339 -----ArgThrValGlnProYrValProLeu.347
DB 1908 AGCGGATGCCGATCTACGAGTGGGTATGCGGATCTCCATCTCGAGACCGCGCTCGAC 1967
QY 348 TrpLeuGlnCysGlnAlaGlyPheTyrValAspAlaGlyPro----- 361
DB 1968 TGG-----TGCATGTCATCAAGTCTGCGAGCACTCGCGGTGGCGGAGCGCTCA 2021
QY 362 ProThrH1sAlaPheThrProCysGlnYlnH1sValCysSerGlnYlnSerAlaYrYrTrp 381
DB 2022 CCGGTCCACCGCGCTGCGCTGCGCGCGCGCGCGCATCGCGAGCACTCGAGCGGATCG 2081
QY 382 SerGlnH1leProLeuProH1sGlyThrH1sAlaPheH1sAlaCys-----Pro 398
DB 2082 CCAAG-----CCTGGCCAGCGCGCGCGCGCTGCGGACATCGGCTGTAGCCAGCGCCC 2138
QY 399 PheCysAlaThrGlnLeuValGlnGln 408
DB 2139 GCCCTCGCGCGCTCGCTCCCGGCGAGCGG 2168

RESULT 14
US-09-949-016-12345/c
; Sequence 12345, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C1001307
; CURRENT APPLICATION NUMBER: US/09/949,016
; PRIOR FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: PaacSeq for Windows Version 4.0
; SEQ ID NO 12345
; LENGTH: 86414
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature

```

```

; LOCATION: (1)...(86414)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-12345

Alignment Scores:
Pred. No.: 147
Score: 107.00
Percent Similarity: 33.3%
Best Local Similarity: 21.9%
Query Match: 4.7%
DB: 3
Gaps: 26

US-10-041-030-4 (1-420) x US-09-949-016-12345 (1-86414)
QY 4 ProGlnGlnGlnH1sCysAlaProAsnYrGlnProValYrGlnYrGlnLeuVal 23
DB 33554 CCGGCGCAGACCGCCACCTGTCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 33516
QY 24 ValLeuGlnYrYrAsnGlnAlaLeuProAsnYrYrAsnGlnYrYrArgYrYrSerArgPhe 43
DB 33515 -----CAAGAGGCTCAGCTCCCTAAGC----- 33492
QY 44 AlaLeuYrYrYrArgPro--LysAlaAsnGlnYrYrProSerThrYrYrH1sVal1le 62
DB 33491 AGCAGCTACCGCTGCGCAGAAAGTACAGGATTTCCAGAGCAACCACTAATGAA 33432
QY 63 SerThrProGlnAla-SerYrYrH1leSerCysYrYrGlnH1sSerTleSerYrTh 82
DB 33431 GTGTGTTTAAAGCTGGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCGCG 33372
QY 82 rleuSerArgAsnGlnThrValValGlnYrYrH1sAspYrYr----- 97
DB 33371 TCTCTCTCACACACACACACACACACACACACACACACACACACACACACACAC 33312
QY 98 -----AspThrAspMetPheGln-- 103
DB 33311 AGTATGATCTGAGCCACTTATGCTGCTGTAAGCTGGAGAGGAGTCTATTAGGC 33252
QY 104 -----ValGlnYrGlnSerThrGlnSerProIleAspPheValVa 116
DB 33251 CTGATCATCAGCGCAGAGGCGATGGGTGGAGAGCGTAAAGGTGGTATACCTTC 33192
QY 116 ThrAspThrTleSerGlnSerGlnAsnThrAspGlnAlaGlnTleThrGlnSerThr 136
DB 33191 GCTGCTCTCCCTGAGTCTGCTCGGAGAGAGAGTCTCTCTCTCAATCTCTCTCTC 33132
QY 136 eSerArgPheAlaCysArgTleValCysAspArgAsnGlnProYrYrTh-AlaArgIleP 156
DB 33131 CAGACATGCTCTGCGCGCTC-----TGCATCTCTCTGCGCGCGCTCT 33087
QY 156 heAlaAlaGlyPheAspSerSerYrYrAsnIlePheLeuGlnYrYrAlaAlaYrYr-- 174
DB 33086 GCGGCGGTGGGAGACATCAGACCGGAG-----CTGGAGCCCAATACACCGAGAGCG 33033
QY 175 -----TrpYrAsnProAspGlnH1s 182
DB 33032 CACCCATGCGCTCCAGGACCGCCCAACGAGATGTGTGGAAACACCGCGGAGACA 32973
QY 182 eAspGlyLeuThrThrAsnGlnYrYrValLeuValMetH1sPro-----ArgGlyY 198
DB 32972 TCAGTGGGCTC-----GTTGGGAAATCCAACTCAGTGGCGCTCTCGAAGAGCAAA 32916
QY 198 YrPheThrGlnGlnSer-----GlnProGlnYrYrTrp----- 208
DB 32915 GG-----GATGTCAGCTTCCCTGCGAGATCCAGAGCTGGCGAGCTGAGTGGCT 32862
QY 209 -----ArgGlnIleSerY 213
DB 32861 GCCAGGTCCCTCTCACTGTGATGCCATCTGTATGTATGAGAGCACTCAAGACA 32802
QY 213 alCysGlyAspValYrYrThrLeu-----ArgGlnThr-----ArgS 225
DB 32801 TCAGCGGCTCTCTCTCAAGTTTGAAGAGGTGTGTCTGACGAGAGACAGAGATCAT 32742

```

```

Qy 225 ealaelnglnlArGlyLys-----LeuValGluSerGluThrAsnVal----- 239
Db 32741 CTCTCTTCGACAGAGGCGACAGAGCTGATGCTGGCACAACCCCTCCAACTCTCTCCAG 32682
Qy 240 -----LeuGlnAerGlySerLeu---LleAerLeuCyGlyValAerThrLeuLeuTrpA 256
Db 32681 GCACAGACTACACACAGAGGCTCCCTGGCCCGCACTTGGCTGGGCGACAGAGCTG----- 32626
Qy 256 rghrhlAaerGlyLeuPheHieThrProThrGlnLysHieGluValAerAerG----- 274
Db 32625 -----GTCACCTCCAAAGCCAGGAATGAGAAACACAGAAAGCTCA 32586
Qy 275 -----GlnGluLleAsnAlaAerProGlnLysProValGlyLeuAerThr- 290
Db 32585 AGCCGAAAGGCCACAGCCACCAACAGCAGCGCGCGGAGACCCAGACCCCGCATTCAC 32526
Qy 291 -----LeuAlaPheProSerLleAerAerGlyGluValAerGluLysGlnP 307
Db 32525 CCAAGCTGAGCAGCAGCCCGACAGCTGACAGCAAGAGCTGAGCTTCAGAAAGGCGAGAC 32466
Qy 307 rorhrlAerGlyLeuSerCyGlyHieValHieGlyTrpHieAerTrpGlyHieAerSerA 327
Db 32465 AGCCCGAGCTCTGAGCATGTGGGGCCCTGGGGCCCTGAGGCTGGGGT----- 32416
Qy 327 ePhrGluAlaAerGluAerGlyLysCyAerProMetCyAerGlyThrValGlyProTrpValAerProL 347
Db 32415 -----GAGCCCCAGATGCGACACCAAGCAGGCTG-----TTGTGCCA- 32380
Qy 347 euTrP-----LeuGlyCyGlyLysGlyPheTrpValAerAerAerAer 361
Db 32379 --TGGGGTCTCGGGTCCAGCTTGGACTGTAAAGGGGTTCT--TAGGACCCCATTC 32325
Qy 361 rorProThrHieAerPheThrProCyGlyHieValCySerGlyLysSerAerAer 379
Db 32324 CAGCTCAGACAGCCAGAAACATGGGTTACAGTGGAGCCAGCAAGGCAAGAG 32269

RESULT 15
US-09-949-016-15758/c
; Sequence 15758, Application US/09949016
; Patent No. 6812339
; GENERAL INFORMATION:
; APPLICANT: VENTER, J. Craig et al.
; TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
; WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
; FILE REFERENCE: C0001307
; CURRENT APPLICATION NUMBER: US/09/949, 016
; CURRENT FILING DATE: 2000-04-14
; PRIOR APPLICATION NUMBER: 60/241,755
; PRIOR FILING DATE: 2000-10-20
; PRIOR APPLICATION NUMBER: 60/237,768
; PRIOR FILING DATE: 2000-10-03
; PRIOR APPLICATION NUMBER: 60/231,498
; PRIOR FILING DATE: 2000-09-08
; NUMBER OF SEQ ID NOS: 207012
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 15758
; LENGTH: 86414
; TYPE: DNA
; ORGANISM: Human
; FEATURE:
; NAME/KEY: misc_feature
; LOCATION: (1)....(86414)
; OTHER INFORMATION: n = A,T,C or G
US-09-949-016-15758

Alignment Scores:
Score: 147 Length: 86414
Percent Similarity: 107.00 Matches: 105
Best Local Similarity: 33.3% Conservative: 55
Query Match: 21.9% Mismatches: 166
DB: 4.7% Indels: 154
Gaps: 26

```

```

US-10-041-030-4 (1-420) x US-09-949-016-15758 (1-86414)
Qy 4 ProGlyGlnGluGlnHieCyValAerProAerGluProValLysTrpGlyGluLeuVal 23
Db 33554 CCGGCGACAGAGCCAGCTGTGGCCCGCACTCCAGCCCT----- 33516
Qy 24 ValLeuGlyTrpAerGlyAlaLeuProAerGlyAerAerGlyAerAerAerAerAerAer 43
Db 33515 -----CAAGAGGCTCAGCTCCCTTAAGC----- 33492
Qy 44 AlaLeuTrpLysAerPro---LysAlaAerGlyValLysProSerThrValHieValLle 62
Db 33491 AGCAGCTACCCCTGGCCACAGAAATGACAGGCAATTTCCAAAGCAACCAATATGAA 33432
Qy 63 SerThrProGlnAerSerValAerLleSerCyGlyGlyGlnHieSerLleSerTrpH 82
Db 33431 GTTGTTTTAAAGGCTGGCCAAATGTCTCTCTCTCCAAATCTCTCTCTCTCTCTCTCT 33372
Qy 82 rLeuSerAerAerGlnThrValValAerGluTrpTrpHieAerLys----- 97
Db 33371 TCTCTCTTCACACACACACACACACACACACACACACACACACACACACACACACACAC 33312
Qy 98 -----AerThrAerMetPheGln-- 103
Db 33311 AGTATGACTGTGACCCACTTATGCTGGGTAAAGTGGAGAGGAGTGACTATTCAGGC 33252
Qy 104 -----ValGlyAerSerThrGlySerProLleAerPheValAer 116
Db 33251 CTGATGATCAGCAGCAGGAGGCGATGGGTGGGTAGAGCAGTGGAGGTGTATCTTC 33192
Qy 116 LThrAerTrpLleSerGlySerGlnAerThrAerGluAlaGlnLleThrGlnSerThrL 136
Db 33191 GCTGCTCCCTCGAAGTCCGTCGGAAGAGCAGTCTCTCTTATGAGGCGCACT 33132
Qy 136 eSerAerPheAlaCyAerGlyLleValCyAerAerAerGluProTrpTrpAlaAerLle 156
Db 33131 CAGAACATGCTCTGGCGGCTC-----TGGCATCTCCCTCGCCACTCT 33087
Qy 156 heAlaAerGlyPheAerSerSerLysAerLlePheLeuGlyLysValAerAerAer 174
Db 33086 GCGGGGTGGGAGCAGCATCAGCAGCGCAG-----CTGGAGCCCAATACACCCAGCAGC 33033
Qy 175 -----TrpLysAerProAerGlyHieAer 182
Db 33032 CACCACTGCGCTCCAGGAGCCCAACAGATTTGTTGGAAACACCCAGCGCAGACA 32973
Qy 182 eAerPlyLeuThrTrpAerGlyValLeuValMetHieAerPro-----ArgGlyAer 198
Db 32972 TCAGTGGGGTCT--GCTGGGGAATGCCAAATCAGTGGGCCCTTCAGAAAGCAGGAATG 32916
Qy 198 LysPheThrGluGluSer-----GlnProGlyValTrp----- 208
Db 32915 GG-----GAGTCCAGCTTCCCTCCCTCCAGATCCAGGCGCTGGCAGCGTGAAGTGGCT 32862
Qy 209 -----ArgGluLleSerV 213
Db 32861 GCCAGTCCCTCTCACTGATGATCCATCTTATGATGAGACAGGCACTCAGAC 32802
Qy 213 aLysGlyAerValLysTrpTrpLeu-----ArgGluTrp-----ArgS 225
Db 32801 TCAGCGGGTCTCTCTCAGGTTGACAGAGGTGTGTCTGACAGGAGACAGAGATCAT 32742
Qy 225 ealaelnglnlArGlyLys-----LeuValGluSerGluThrAsnVal----- 239
Db 32741 CTCTCTTCGACAGAGGCGACAGAGCTGATGCTGGCACAACCCCTCCAACTCTCTCCAG 32682
Qy 240 -----LeuGlnAerGlySerLeu---LleAerLeuCyGlyValAerThrLeuLeuTrpA 256
Db 32681 GCACAGACTACACACAGAGGCTCCCTGGCCCGCACTTGGCTGGGCGACAGAGCTG----- 32626
Qy 256 rghrhlAaerGlyLeuPheHieThrProThrGlnLysHieGluValAerAerG----- 274

```

```
Db 32625 -----GTCACTCCAGAGCCAGAAAATGAGAACACAGAAAGCTCA 32586
Qy 275 -----GInguIuIlleAsnAlaAlaArgProGInCySProValGlyLeuAsnThr- 290
Db 32585 AGCCGAAGGCCAGCCAGCCAGCAGAGCCGAGTGGAAGCCGACCCGATACAC 32526
Qy 291 -----LeuAlaPheProSerIleAsnArgLySGluValValGluGluLeuGlnP 307
Db 32525 CCAAGCTGCGACAGCAGCCCGACACTGACAGAGAGCTGAGCTTCCAGAAAGGCGAGAC 32466
Qy 307 roTTrpAlaTyrlleuSerCySGlyHisValHisGlyTyrlHisAsnTrpGlyHisArgSera 327
Db 32465 AGCCCGAGCTCTCAGCATGTGGGGGCCCTGGGGCCCTGTGAGGCTGSGGT----- 32416
Qy 327 sptHrGluAlaAngluArgGluCySPrometCySArgThrValGlyProTyValProl 347
Db 32415 -----GAGCCCCAGATGTGCAGACCAAGGC---TTGTCCCA- 32380
Qy 347 euTrp-----LeuGlyCySGluAlaGlyPheTyValAspAlaGlyP 361
Db 32379 --TGGGTCCTGGGTTCCAGCTTTGACTGTMAAATGGGGTTC--TAGAACCCCATTC 32325
Qy 361 roPTrHrHisAlaPheThrProCySGlyHisValCySserGluLySeraLalys 379
Db 32324 CAGCTCAGACAGCCAGAAACATGGGTTCAACAGTGGAGGCCAAGACCAAGAG 32269
```

Search completed: March 2, 2006, 09:43:27
Job time : 3232 secs

this page blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - nucleic search, using frame_p2n model

Run on: March 2, 2006, 05:14:56 ; Search time 5360 Seconds
(without alignments)
3666.152 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MNSPQGEHCAPKPEVKYK.....ATQVGEQNCIKIKIFQGPID 420

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 41078325 seqs, 23393541228 residues
Total number of hits satisfying chosen parameters: 82156650

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-O=/abs/ABSWEB.spool/US10041030/runat 01032006 134406 21863/app query.fasta_1
-DB=EST -OPMT=fastap -SUFFIX=p2n.rst -MINMATCH=0.1 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdi -LIST=45
-DOCALLGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs808
-USER=US10041030 @CGM.1.1 5315 @runat 01032006 134406 21863 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

EST:
1: gb_est1:
2: gb_est2:
3: gb_est3:
4: gb_est4:
5: gb_est5:
6: gb_est6:
7: gb_est7:
8: gb_est8:
9: gb_est9:
10: gb_est10:
11: gb_est11:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|-------------|--------------------|
| 1 | 2151.5 | 94.0 | 3387 | 4 AK030564 | Mus muscu |
| 2 | 2147 | 93.8 | 1186 | 10 AY409116 | AY409116 Homo sapi |
| 3 | 2052.5 | 89.6 | 1183 | 10 AY409118 | AY409118 Mus muscu |
| 4 | 1926 | 84.1 | 3530 | 4 AK045515 | AK045515 Mus muscu |
| 5 | 1912 | 83.5 | 2584 | 4 AK045673 | AK045673 Mus muscu |
| 6 | 1681 | 73.4 | 1186 | 10 AY409117 | AY409117 Pan trogl |
| 7 | 1632 | 71.3 | 1879 | 4 AK044418 | AK044418 Mus muscu |

| | | | | | |
|----|--------|------|------|-------------|----------------------|
| 8 | 1626.5 | 71.0 | 1242 | 10 AY408790 | AY408790 Pan trogl |
| 9 | 1622 | 70.8 | 2572 | 4 CR606152 | CR606152 full-length |
| 10 | 1621.5 | 70.8 | 1242 | 10 AY408789 | AY408789 Homo sapi |
| 11 | 1620 | 70.7 | 2710 | 4 HSM805479 | AL834395 Homo sapi |
| 12 | 1611.5 | 70.4 | 1242 | 10 AY408791 | AY408791 Mus muscu |
| 13 | 1563 | 68.3 | 2474 | 4 AK033815 | AK033815 Mus muscu |
| 14 | 1519 | 66.3 | 849 | 5 BX388547 | BX388547 BX388547 |
| 15 | 1495 | 65.3 | 855 | 8 CX975821 | CX975821 JGI CAAP9 |
| 16 | 1473 | 64.3 | 3531 | 4 AK029886 | AK029886 Mus muscu |
| 17 | 1430.5 | 62.5 | 957 | 7 CN329888 | CN329888 AGENCOURT |
| 18 | 1391 | 60.7 | 869 | 5 BX688136 | BX688136 BX688136 |
| 19 | 1363.5 | 59.5 | 856 | 6 CF241089 | CF241089 AGENCOURT |
| 20 | 1362 | 59.5 | 1117 | 7 CK027873 | CK027873 AGENCOURT |
| 21 | 1352 | 59.0 | 828 | 8 CX314078 | CX314078 JGI XZT96 |
| 22 | 1344.5 | 58.7 | 879 | 8 CX941650 | CX941650 JGI CAAP9 |
| 23 | 1334.5 | 58.3 | 1076 | 3 BM800644 | BM800644 AGENCOURT |
| 24 | 1314 | 57.4 | 809 | 5 BX700555 | BX700555 BX700555 |
| 25 | 1306.5 | 57.1 | 764 | 3 BP682660 | BP682660 BP682660 |
| 26 | 1294 | 56.5 | 701 | 7 CK637810 | CK637810 UI-M-HO0 |
| 27 | 1261 | 55.1 | 768 | 7 CO738704 | CO738704 SLIB04C21 |
| 28 | 1259.5 | 55.0 | 726 | 3 BP701801 | BP701801 BP701801 |
| 29 | 1258 | 54.9 | 784 | 5 BX749714 | BX749714 BX749714 |
| 30 | 1237 | 54.0 | 1229 | 4 BC026695 | BC026695 Mus muscu |
| 31 | 1207 | 52.7 | 741 | 8 CX744549 | CX744549 JGI ANBT2 |
| 32 | 1203.5 | 52.6 | 752 | 3 BP709958 | BP709958 BP709958 |
| 33 | 1192.5 | 52.1 | 1188 | 8 DN729302 | DN729302 CMB46-E04 |
| 34 | 1189 | 51.9 | 747 | 5 BX774759 | BX774759 BX774759 |
| 35 | 1187.5 | 51.9 | 773 | 8 CX439604 | CX439604 JGI XZG77 |
| 36 | 1187 | 51.8 | 738 | 8 CX941649 | CX941649 JGI CAAP9 |
| 37 | 1187 | 51.8 | 752 | 8 CX443252 | CX443252 JGI XZG89 |
| 38 | 1185 | 51.7 | 1006 | 3 BO049269 | BO049269 AGENCOURT |
| 39 | 1183 | 51.7 | 806 | 7 CR975985 | CR975985 CR975985 |
| 40 | 1169 | 51.0 | 732 | 8 CX436008 | CX436008 JGI XZG58 |
| 41 | 1161 | 50.7 | 766 | 6 CF539872 | CF539872 UI-M-BX0 |
| 42 | 1159 | 50.6 | 889 | 9 BX327874 | BX327874 BX327874 |
| 43 | 1155.5 | 50.5 | 832 | 8 DN521048 | DN521048 1263866 M |
| 44 | 1154.5 | 50.4 | 795 | 7 CK638809 | CK638809 UI-M-HO0 |
| 45 | 1148 | 50.1 | 763 | 5 BM959532 | BM959532 BM959532 |

ALIGNMENTS

| | | | | | |
|------------|--|-------------|------|--------|-----------------|
| RESULT 1 | AK030564 | 3387 bp | mRNA | linear | HTC 03-Apr-2004 |
| LOCUS | AK030564 | | | | |
| DEFINITION | Mus musculus adult male pituitary gland cDNA, RIKEN full-length enriched library, clone:530432G13 product:pellino 2, full insert sequence. | | | | |
| ACCESSION | AK030564 | GI:26326560 | | | |
| VERSION | AK030564.1 | | | | |
| KEYWORDS | HTC; CAP trapper. | | | | |
| SOURCE | Mus musculus (house mouse) | | | | |
| ORGANISM | Mus musculus | | | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | | |
| 1 | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komuro, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. | | | | |
| 2 | High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999) | | | | |
| 10349636 | | | | | |
| 2 | Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K., Itoh, M., Komuro, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y. | | | | |
| 3 | Normalisation and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000) | | | | |
| 11042159 | | | | | |
| 3 | Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P., Komuro, H., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., | | | | |

Fujiwaka, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watanabe, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsumura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system—384-format sequencing pipeline with 384 multiplexed capillary sequencer Genome Res. 10 (11), 1757-1771 (2000)

4 The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

5 The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

6 (bases 1 to 3387) Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hirooka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishii, K., Nomura, K., Numazaki, R., Ohno, M., Ohsato, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Tanaka, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M., and Hayashizaki, Y. Direct Submission

Submitted (16-JUL-2001) Yoshihide Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute; 1-7-22 Suehiro-cho, Tsurumi-Ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

CDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

1. .3387

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="FANTOM_DB:5330432G13"

/db_xref="taxon:10090"

/clone="5330432G13"

/sex="male"

/tissue_type="pituitary gland"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

336. .1595

/note="unnamed protein product; pellino 2 (MGI:1891445, GB|NM_033602, evidence: BLASTN, 99%, match=864)

putative"

/codon_start=1

/protein_id="BAC27024.1"

/db_xref="GI:26326561"

/translation="MSPSGKPEPTPKVKKYRELIVYGYNALPVGDRGKSRPA LYKRTASGVKPESTIHNVSTPQASKAISGSHSISTTSSRSQTVVEYTHDQTFK QVNRSTSLIDFVDTVSGQNDADQITSTISRPFGCTVLCRNEEYTKRIADAGD SSKNFIKGEKAKMKNPDGMDGLTNGVIVMHPGFTESQPGVREISVCGDVT LRETRSGKRGKVESRTNLDGSLIDLGAATLWRTADGLFPAPTOKHIALROBI NNAAPQCPVGLNTLAPPSINRKEVVEKQGMAYLSCGHVGHYSWGRSDTEANREBC

ORIGIN

Alignment Scores:

Pred. No.: 1,92e-205 Length: 3387

Score: 2151.50 Matches: 395

Percent Similarity: 96.7% Conservative: 11

Best Local Similarity: 94.0% Mismatches: 13

Query Match: 94.0% Indels: 1

DB: 4 Gaps: 1

US-10-041-030-4 (1-420) x AK030564 (1-3387)

Qy 1 MetPheSerProGlyInGluGluHisCysAlaProAsnGlyGluProValIysTrgIy 20

Db 336 ATTTTTCCTCCCGGGCCANAAAGAACCCANCAACCCCAAGAGANCCGGTAAATACCGG 395

Qy 21 GluLeuValValIleuGlyTrAengIyAlaLeuProAsnGlyAspArgGlyArgArgIys 40

Db 396 GAGCTGGTGCTCTGGGTRCAATGCTGCTTACCTAATGCTGACAGGGGAGAGGANA 455

Qy 41 SerArgPheAlaLeuTrIyIysArgProIyAlaAsnGlyValIysProSerThrValHis 60

Db 456 AGCAGATTTGGCCCTTAAAGCGGACCTACGCCAGTGGTGCMAACCCAGCAATCCAC 515

Qy 61 ValIleSerThrProGlnIaSerIyAlaIleSerCysIyGlyGluHisSerIleSer 80

Db 516 ATGCTTCCACACACAGCGGCTCCAGGATGCTCCAGAGACATCCAGCATATCG 575

Qy 81 TrThrIleuSerArgAsnGlnThrValValIleGluTrThrIleAspIysAspThrAsp 100

Db 576 TACAGTTGTCACGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGC 635

Qy 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValIleThrAspThrIle 120

Db 636 ATGTTTCAGTGGGCGAGTCAACAGAAAGCCCATTTGACTTCTGCTGCACAGCAGGTT 695

Qy 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140

Db 696 TCGGCGGTCAGAACAGAAAT--GCCAGATCACAGAGCAGCAGCAGCAGCAGCAGCAGCAG 752

Qy 141 CysArgIleValCysAspArgAsnGluProIyTrThrAlaArgIlePheAlaIleArgIle 160

Db 753 TGCAGATCTGCTGTACAGAAAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 812

Qy 161 AspSerSerIyAsnIlePheLeuGlyGluIyAlaIleIyAspIyAsnProAspGly 180

Db 813 GATTTCTCCAAAATATCTTTCTTGGAGAGAAAGCAGAAATGGAATAAACCCTGATGGA 872

Qy 181 HisMetAspGlyLeuThrThrAsnGlyValIleValMetHisProArgGlyGlyPheThr 200

Db 873 CACATGATGATGATCTACCAACAGGCTGCTGATGATGACCCGAGAGAGGCTTCACC 932

Qy 201 GluGlySerGlnProGlyValIleTPAArgIuIleSerValCysGlyAspValIyTrThrIleu 220

Db 933 GAGGAATCCAGGCTGAGATCTGAGAGAGATCTCTGCTGAGGATGATACACTTG 992

Qy 221 ArgGluThrArgSerAlaGlnGlnArgGlyIyLeuValGluSerGluThrAsnValIleu 240

Db 993 CGAGAACACAGGTCGAGCCGAG 1052

Qy 241 GlnAspGlySerIleuIleAspLeuCyGlyAlaIleThrIleuLeuTrpArgThrAlaAspGly 260

Db 1053 CAAAGCGGCTCCCTCATTTGACCTGTGTGGGCCACTCTCTCGAGAAACCGAGATGGC 1112

Qy 261 LeuPheHisThrProThrGlnIyHisIleGluAlaIleuArgGlnIuIleAsnAlaIle 280

Db 1113 CTTTTCACGCTCTTACTCAGAGCAGCAGAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAGCAG 1172

Qy 281 ArgProGlnCysProValGlyLeuAsnThrIleuAlaPheProSerIleAsnArgIyGlu 300

Db 1173 CGACCCAGTGGCCCGTGGGCTTAAACCTGCTTCCCGCAGATCAACCGAGAGAA 1232

| | | | | | |
|------------------------|---|---|----------------------------|--|--|
| Qy | 301 | ValValGluGluGluValGluGluGluProTThrAlaGlyTyrLeuSerGlyGlyIshValIshGlyTyrHis | 320 | | |
| Db | 1233 | GTGGTGTGAAGAAGAACGACCCCTGGGCATATCCTGAGCTGGCCCATGTGCACGCGCTACAC | 1292 | | |
| Qy | 321 | AaaTrrGlyIshIshASGSerAspTrrGluIshIshIshValIshIshGluGluCysProMetCysArgThr | 340 | | |
| Db | 1293 | AGCTGGGGCCATCGGAGGGAACAGGAACCAACAGAGGAGTGTCCCATGTGCAGACT | 1352 | | |
| Qy | 341 | ValGlyProTyrValProLeuTrrPleuGlyCysGluAlaGlyPheTyrValAspAlaGly | 360 | | |
| Db | 1353 | GTGGGCGCCCTACGTCCTCTCTGCTGGGCTGTGAGGACGAGTATTTATGTGCAGCCGGA | 1412 | | |
| Qy | 361 | ProProThrIshIshAlaPheThrProCysGlyIshIshValCysSerGluValSerAlaIshTyr | 380 | | |
| Db | 1413 | CCCCCAATCAACGCTTTCACCCCTGGGGGACGCTGTGTAGAGAACTGTCCAGATAC | 1472 | | |
| Qy | 381 | TrrSerGluIshProLeuProIshGlyTrrIshIshAlaPheIshIshAlaIshCysProPheCys | 400 | | |
| Db | 1473 | TGGTGTGAGATCCCACTGCCCCACAGGAACGACGCGTTTTCACGCGCGCTGTCCGTTCTGC | 1532 | | |
| Qy | 401 | AlaThrGluLeuValGlyGluGluIshIshCysIshIshLeuIshIshPheGlnGlyProIshAsp | 420 | | |
| Db | 1533 | GCCACGCGACGCTGTGGTGTGAACAGAACTGCATCAATGATTATTCAGAGGTCCAGGTGAC | 1592 | | |
| RESULT 2 | | | | | |
| LOCUS | AY409116 | 1186 bp | DNA linear GSS 16-DEC-2003 | | |
| DEFINITION | Homo sapiens PEL12 gene, VIRUAL TRANSCRIPT, partial sequence, | | | | |
| ACCESSION | AY409116 | genomic survey sequence. | | | |
| VERSION | AY409116.1 | GI:39765084 | | | |
| KEYWORDS | GSS. | | | | |
| SOURCE | Homo sapiens (human) | | | | |
| ORGANISM | Homo sapiens | | | | |
| | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | |
| | Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; | | | | |
| | Hominae; Homo. | | | | |
| REFERENCE | 1 (bases 1 to 1186) | | | | |
| AUTHORS | Clark,R.A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., | | | | |
| | Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., | | | | |
| | Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., | | | | |
| | Adams,M.D. and Cargill,M. | | | | |
| TITLE | Inferring nonneutral evolution from human-chimp-mouse orthologous | | | | |
| | gene trios | | | | |
| JOURNAL | Science | 302 (5652), | 1960-1963 (2003) | | |
| PUBMED | 14671302 | | | | |
| REFERENCE | 2 (bases 1 to 1186) | | | | |
| AUTHORS | Clark,R.A.G., Glanowski,S., Nielson,R., Thomas,P., Kejarival,A., | | | | |
| | Todd,M.A., Tenenbaum,D.M., Civello,D.R., Lu,F., Murphy,B., | | | | |
| | Ferreira,S., Wang,G., Zheng,X.H., White,T.J., Shinsky,J.J., | | | | |
| | Adams,M.D. and Cargill,M. | | | | |
| TITLE | Direct Submission | | | | |
| JOURNAL | Submitted (16-NOV-2003) | Celera Genomics, 45 West Gude Drive, | | | |
| | Rockville, MD 20850, USA | | | | |
| COMMENT | This sequence was made by sequencing genomic exons and ordering | | | | |
| | them based on alignment. | | | | |
| FEATURES | Location/Qualifiers | | | | |
| source | 1..1186 | | | | |
| | /organism="Homo sapiens" | | | | |
| | /mol_type="genomic DNA" | | | | |
| gene | /db_xref="taxon:9606" | | | | |
| | <1..>1186 | | | | |
| | /gene="PEL12" | | | | |
| | /locus_tag="HCM3450" | | | | |
| ORIGIN | | | | | |
| Alignment Scores: | | | | | |
| Pred. No.: | 1,1e-205 | Length: | 1186 | | |
| Score: | 2147.00 | Matches: | 394 | | |
| Percent Similarity: | 100.0% | Conservative: | 0 | | |
| Best Local Similarity: | 100.0% | Mismatches: | 0 | | |
| Query Match: | 93.8% | Indels: | 0 | | |

| DB: | 10 | Gaps: | 0 |
|---|-----|---|------|
| US-10-041-030-4 (1-420) x AY409116 (1-1186) | | | |
| QY | 27 | TyRANGIYAlAlaLeuProANGIYAspArgIYArgArgLYSerArgPheAlaLeuTYr | 46 |
| Db | 2 | TRCAATGGTGGTTTACCCAAATGAGATATGAGACGAGGAAAGTAAGATTGGCTTAC | 61 |
| QY | 47 | LYsArgProLYsAlaANGIYAlaLYsProSerThrValHisValIleSerThrProGln | 66 |
| Db | 62 | AAGCGGCCCAAGCGAAATGTGTCAAAACCCAGCAACCTTCATGTGATATCCACGCCCAAG | 121 |
| QY | 67 | AlaSerLYsAlaIleSerCYsLYsGLYGLHisSerIleSerTYrThrLeuSerArgAsn | 86 |
| Db | 122 | GCATTCAAAGGCTATCAGCTGCAAAAGGTCAACACAGTATATCTTACCTTGTCAAGAAAT | 181 |
| QY | 87 | GlnThrValAlaValGluTYrThrHisAspLYsAspThrAspMetPheGlnValGlyArg | 106 |
| Db | 182 | CAGACTGTGTGTGTGAGATACACATATATAGAGATATAGATTATTTCAAGTGGCAGA | 241 |
| QY | 107 | SerThrGluSerProIleAspPheValValIleAspThrIleSerGlySerGlnAsnThr | 126 |
| Db | 242 | TCAACAGAAACCCCTATGACTGTGTGTCTCAGACACAGATTCTTGGGACGACAGAACG | 301 |
| QY | 127 | AspGluAlaGlnIleThrGlnSerThrIleSerArgPheAlaCYsArgIleValCYsAsp | 146 |
| Db | 302 | GACGAAAGCCAGATACACAGAGCAACCATATCCAGTTCGCTGAGAGATGTGTGCAC | 361 |
| QY | 147 | ArgANGIuProTYrThrAlaArgIlePheAlaAlaGlyPheAspSerSerLYsAsnIle | 166 |
| Db | 352 | AGGATGAACCTTACACGACAGATATTCGCCCGGATTTGACTCTTCCAAATAACTA | 421 |
| QY | 167 | PheLeuGlyGluLYsAlaAlaLYsTrpLYsAsnProAspGlyHisMetAspGlyLeuThr | 186 |
| Db | 422 | TTCTTGGAGAAAGGACAGCAAAAGTGAATAAACCCGACGCGCATGATGGGCTCACT | 481 |
| QY | 187 | ThrANGIYAlaLeuValMetHisProArgGlyGlyIlePheThrGluGluSerGlnProGly | 206 |
| Db | 482 | ACTAATAGCGCTCGTGATGATCATCCAGAGGGGCTTCAACGAGATCCACGCCCGAG | 541 |
| QY | 207 | ValTrpArgGluIleSerValCYsGlyAspValTYrThrLeuArgGluThrArgSerAla | 226 |
| Db | 542 | GTTCTGGCGAGATTTCTGTCTGTGAGATGTGTACCTTGGAGAAACAGATCGGCC | 601 |
| QY | 227 | GlnGlnArgGlyLYsLeuValGluSerGluThrAsnValLeuGlnAspGlySerLeuIle | 246 |
| Db | 602 | CAGCAAGAGAAAGCTGTGGAAGTGAACCAACGTCCTGACAGACGGCTCCCTCATTT | 661 |
| QY | 247 | AspLeuCYsGlyAlaThrLeuLeuTrpArgThrAlaAspGlyLeuPheHisThrProThr | 266 |
| Db | 652 | GACCTGTGTGGGGCCACTCTCTCTCGAGAAACAGCAGATGGGCTTTTTCATATCTCAACT | 721 |
| QY | 267 | GlnLYsHisGlnAlaLeuArgGlnGlnIleAsnAlaAlaArgProGlnCYsProVal | 286 |
| Db | 722 | CAGAAAGCAATAGAAAGCCCTCGAGAGAAATTAACGCCGCCGCTCAAGTCTCTGTG | 781 |
| QY | 287 | GlyLeuAsnThrLeuAlaPheProSerIleAsnArgLYsGluValAlaGluGluLYsGln | 306 |
| Db | 782 | GGGGCTCAACACCTGTGGCTTCCCAAGCATCAACAGAAAGAGTGTGTGAGAGAAAGAG | 841 |
| QY | 307 | ProTrpAlaTYrLeuSerCYsGlyHisValHisGlyTYrHisAsnTrpGlyHisArgSer | 326 |
| Db | 842 | CCCTGGGAGATATCTCATGTTGTGGCCACAGTCGACCGGGTACCAACCTGGGGCCATGAGAT | 901 |
| QY | 327 | AspThrGluAlaAsnGlnAlaGluCYsArgMetSerValGluThrValGlyProTYrValPro | 346 |
| Db | 902 | GACAGAGAGGCCAAGAGAGAGAGTGTCCATGTGACAGACTGTGTGGGCCCTATGTGCTT | 961 |
| QY | 347 | LeuTrpLeuGlyCYsGlyAlaGlyPheTYrValAspAlaGlyProProThrHisAlaPhe | 366 |
| Db | 962 | CTCTGGCTTGGGTGTGAGGAGAGATTTATGTAGACGAGAGCCGCAACTCATGCTTTTC | 1022 |
| QY | 367 | ThrProCYsGlyHisValCYsSerGlyLYsSerAlaLYsTYrTrpSerGlnIleProLeu | 386 |

Db 1022 ACTCCCTGGAGACGCTGCTCGAGAGAGCTGCAAAATATGCTCAGATCCCGTTG 1081
Qy 387 ProHieGlyThrhIaIaPheHieIaIaCyseProPheCyseIaIaThrhInIeUaIaGly 406
Db 1082 CCTCATGGAACCTCATGATTTTCACGCTGCTGCTTCTGTGTACACAGCTGGTGGG 1141
Qy 407 GlUGlnAsnCysIleUaIeUaIePheGlnGlyProIleAsp 420
Db 1142 GAGGAAACCTGCATCAATTAATTTTCCAGATCCATTCAC 1183
RESULT 3
AY409118 1183 bp DNA linear GSS 16-DEC-2003
LOCUS Mus musculus PELI2 gene, VIRUAL TRANSCRIPT, partial sequence, 1081
DEFINITION genomic survey sequence.
ACCESSION AY409118 GI:39765086
VERSION AY409118.1
KEYWORDS GSS.
SOURCE Mus musculus (house mouse)
ORGANISM Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
Sciurognathi; Muridae; Murinae; Mus.
REFERENCE 1 (bases 1 to 1183)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Smitsky, J.J.,
Adams, M.D. and Cargill, M.
Interferring nonneutral evolution from human-chimp-mouse orthologous
gene trios
JOURNAL Science 302 (5652), 1960-1963 (2003)
PUBMED 14671302
AUTHORS 2 (bases 1 to 1183)
Clark, A.G., Glanowski, S., Nielson, R., Thomas, P., Kejarival, A.,
Todd, M.A., Tanenbaum, D.M., Civeello, D.R., Lu, F., Murphy, B.,
Ferreira, S., Wang, G., Zheng, X.H., White, T.J., Smitsky, J.J.,
Adams, M.D. and Cargill, M.
Direct Submission
Submitted (16-NOV-2003) Celera Genomics, 45 West Gude Drive,
Rockville, MD 20850, USA
COMMENT This sequence was made by sequencing genomic exons and ordering
them based on alignment.
FEATURES
source location/Qualifiers
1..1183
/organism="Mus musculus"
/mol_type="genomic DNA"
/db_xref="taxon:10090"
<1..>1183
/gene="PELI2"
/locus_tag="HGM3450"
ORIGIN
Alignment Scores:
Pred. No.: 3.84e-196 Length: 1183
Score: 2052.50 Matches: 375
Percent Similarity: 97.7% Conservative: 10
Best Local Similarity: 95.2% Mismatches: 8
Query Match: 89.6% Indels: 1
Gaps: 1
US-10-041-030-4 (1-420) x AY409118 (1-1183)
Qy 27 TyraHengIyAlaIeuProAengIyAspArgIyAraGlyIySeSerArgPheIaIeUaIy 46
Db 2 TyraHengIyAlaIeuProAengIyAspArgIyAraGlyIySeSerArgPheIaIeUaIy 46
Qy 47 TyraHengIyAlaIeuProAengIyAspArgIyAraGlyIySeSerArgPheIaIeUaIy 46
Db 62 TyraHengIyAlaIeuProAengIyAspArgIyAraGlyIySeSerArgPheIaIeUaIy 46
Qy 67 TyraHengIyAlaIeuProAengIyAspArgIyAraGlyIySeSerArgPheIaIeUaIy 46
Db 121 TyraHengIyAlaIeuProAengIyAspArgIyAraGlyIySeSerArgPheIaIeUaIy 46

Db 122 GCCTCCAGAGCCATCAGCTCCAGAGACATCAGACATATCTGACCTGTTCACGAGC 181
Qy 87 GlnThValIa 106
Db 182 CAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 211
Qy 107 SerThrGlnSerProIleAspPheValIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 126
Db 242 TCACAGGAAAGCCCATTTGATCTTCTGTGTACAGACAGGCTTCCGCGGTTCAGAGAA 301
Qy 127 AspGlnIa 146
Db 302 GAT---GCCAGATACACAGACACACACACACACACACACACACACACACACACACAC 358
Qy 147 ArgAsnGlnProIyThrhIa 166
Db 359 AGAAGAGCCATATACAGAC 418
Qy 167 PheIaGlyIy 186
Db 419 TTTCTTGAGAGAAAGCAGCAAAATGAAATGAAATGAAATGAAATGAAATGAAATGAAAT 478
Qy 187 ThraHengIyAlaIeUaIeUaIeUaIeUaIeUaIeUaIeUaIeUaIeUaIeUaIeUa 206
Db 479 ACCAATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 538
Qy 207 ValThraArgGlnIleSerValIyAspValIyThraArgGlnIleSerValIyAspVal 226
Db 539 GTCTGAGAGAGATCTCTGT 598
Qy 227 GlnGlnIa 246
Db 599 CAGCAGAGGAGAAAGCTGT 658
Qy 247 AspLeuCyseIyAlaIa 266
Db 659 GACCTGT 718
Qy 267 GlnIyHieIa 286
Db 719 CAGACGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 778
Qy 287 GlnIyHieIa 306
Db 779 GGCCTTAAACCTTGGCTTCCACATCAACCGAAGAGAGAGAGAGAGAGAGAGAGAGAG 838
Qy 307 ProIyAlaIy 326
Db 839 CCCTGGGCAATACCTGAGCTGGGCAATGAGCAAGCTGAGCAAGCTGAGCAAGCTGAG 898
Qy 327 AspThrhIa 346
Db 899 GAGCGAG 958
Qy 347 LeuThrpLeuGlyIyCyseGlnIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 366
Db 959 CTCTGCTGGGCTGTGAGAGAGATTTTATGATGATGATGATGATGATGATGATGATGAT 1018
Qy 367 ThrpProCyseIy 386
Db 1019 ACCCCCTGGGAG 1078
Qy 387 ProHieGlyThrhIa 406
Db 1079 CCCACAG 1138
Qy 407 GlUGlnAsnCysIleUaIeUaIePheGlnGlyProIleAsp 420
Db 1139 GAGGAAACCTGCATCAATTAATTTTCCAGATCCATTCAC 1180
RESULT 4
AK045515 3530 bp mRNA linear HTC 03-APR-2004

DEFINITION Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length enriched library, clone:B230208C14 product:pellino 2, full insert sequence.

ACCESSION AK045515

VERSION AK045515.1

KEYWORDS HTC; CAP trapper.

SOURCE Mus musculus (house mouse)

ORGANISM Mus musculus

REFERENCE Carninci, P. and Hayashizaki, Y. High-efficiency full-length cDNA cloning Meth. Enzymol. 303, 19-44 (1999)

REFERENCE Carninci, P. and Hayashizaki, Y. Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes Genome Res. 10 (10), 1617-1630 (2000)

REFERENCE Shibata, K., Itoh, M., Aizawa, K., Nagasaka, S., Sasaki, N., Carninci, P., Kono, H., Akiyama, J., Nishi, K., Kikunai, T., Tashiro, H., Itoh, M., Sumi, N., Ishii, Y., Nakamura, S., Hazama, M., Nishine, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujisawa, S., Inoue, K., Togawa, Y., Ozawa, K., Ohara, E., Matshiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsunaga, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A. and Hayashizaki, Y. RIKEN integrated sequence analysis (RISA) system-384-format sequencing pipeline with 384 multipillar sequencer Genome Res. 10 (11), 1757-1771 (2000)

REFERENCE The RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium. Functional annotation of a full-length mouse cDNA collection Nature 409, 685-690 (2001)

REFERENCE The PANTOM Consortium and the RIKEN Genome Exploration Research Group Phase I & II Team. Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs Nature 420, 563-573 (2002)

REFERENCE Adachi, J., Aizawa, K., Akimura, T., Arahata, T., Bono, H., Carninci, P., Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W., Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T., Horii, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T., Kato, H., Kawai, J., Kojima, Y., Kondo, S., Kono, H., Kouda, M., Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Murata, M., Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N., Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T., Sogabe, Y., Tagami, M., Tagawa, A., Takahashi, F., Takaku-Akahira, S., Takeda, Y., Tanaka, T., Tomaru, A., Toya, T., Yasunishi, A., Muramatsu, M. and Hayashizaki, Y. Direct Submission

TITLE Submitted (16-JUL-2001) Yoshihide Hayashizaki. The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), RIKEN Yokohama Institute, 1-7-22 Suehiro-cho, Tsurumi-ku, Yokohama, Kanagawa 230-0045, Japan (E-mail: genome-res@gscc.riken.jp, url: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216)

COMMENT cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in RIKEN. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues.

Please visit our web site for further details.

FEATURES

source

URL: http://genome.gsc.riken.jp/
URL: http://fantom.gsc.riken.jp/
Location/Qualifiers

1..3530

/organism="Mus musculus"

/mol_type="mRNA"

/strain="C57BL/6J"

/db_xref="PANTOM:DB:B230208C14"

/db_xref="taxon:10090"

/clone="B230208C14"

/sex="male"

/tissue_type="corpora quadrigemina"

/clone_lib="RIKEN full-length enriched mouse cDNA library"

/dev_stage="adult"

292..1421

/note="pellino 2 (MGD|WGI:1891445, GB|NM_033602, evidence: BLASTN, 99%, match=864)

3513..3518

/note="putative"

3530

/note="putative"

polyA_site

3530

polyA_site

3530

/note="putative"

ORIGIN

Alignment Scores:

Pred. No.: 1.23e-182 Length: 3530

Score: 1926.00 Matches: 361

Percent Similarity: 87.4% Conservative: 6

Best Local Similarity: 86.0% Mismatches: 9

Query Match: 84.1% Indels: 45

DB: 4 Gaps: 2

US-10-041-030-4 (1-420) x AK045515 (1-3530)

1 MetPhaserProGlyngingluGluH1sCyAlaProAsnlysgLuproVallystyGly 20

292 ATGTTTCCCGGGCGAGAGGAAACCAGCGCCCAACAGAGCGCGTGAATACGGG 351

21 GluLeuValValleuGlyTyraenGlyAlaLeuProAsnlyAspArgGlyArgGly 40

352 GACCTGGTGCTCT-GGG----- 368

41 SerArgPheAlaLeuTyrylsArgProlyAlaAsnGlyVallysrProserThrValHis 60

368 ----- 368

61 ValIleSerThrProGlnAlaSerlyAlaIleSerCyelysgLyGlnHisSerIleSer 80

369 -----GCCATACGCTCCAGAGGACATCCACATATCG 401

81 TyrThrLeuSerArgAsnGlnThrValValleuGlyTyrrHisAspLyAspThrAsp 100

402 TACACGTTCACAGGAGCGAGCGATAGTGTGAGTACACACACATTAAGACACAGAC 461

101 MetPheGlnValGlyArgSerThrGlySerProIleAspPheValValThrAspThrIle 120

462 ATGTTTCAGGTGGCGAGGTTCACAGAAAGCCCATGACTCTGTGTTCACAGACGGT 521

121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140

522 TCCGGCGGTGCAACGAAAGT---GCCCATACAGGACGACCATCTTACGTTCCGA 578

141 CysArgIleValCysAspArgAsnGluProTyrrThrAlaArgIlePheAlaArgIlePhe 160

579 TGCAGAGTGTGTGTGACAGGAACGACATACAGACGCAATATTCGGCAGAGATTC 638

161 AspSerSerlyAsnIlePheLeuGlyGluValAlaAlaIleTyrrlyAsnProAspGly 180

639 GATTTCTCCAAAATATCTTTCTTGGAGGAAAGCAGAAATGAAAAAATCCCTGATGGA 698

181 HisMetAspGlyLeuThrThrAsnGlyValleuValMetHisProArgGlyGlyPheThr 200

699 CACATGATGACTCTACATACCAACGCTGTCTGTATGACACCCCGCAGAGAGCTTCACC 758

| | | | | |
|-----------------|---|------|--|-----------------------------|
| Oy | | 201 | GluGusSerGlnProGluValTPrpArgLuuIleSerValCySGLVASpValTyThrLeu | 220 |
| Db | | 759 | GAGCAATCCACACCCTCGAGACTCGAGAGAAGATCTCTGTCTGGGGAACTATACTTC | 818 |
| Oy | | 221 | ArgGluThraGSeSerAlaGInGlnAargLyuLyLeuValGluSeGluThraAsnValIeu | 240 |
| Db | | 819 | CGAGAGAACCAAGTCCGCCACAGCAGGGGAAAAGCTGTGAAGAGAACCAACGTCCTG | 878 |
| Oy | | 241 | GlnAspGlySeSerIleuIlaAspLeucyGylValaThrLeuleuThraXgThraIlaAspLyu | 260 |
| Db | | 879 | CNAGCGGCCTCCCTATTGA CCTGTGTGGGGCACATCTCTCTTGAGATCGCAATATGC | 938 |
| Oy | | 261 | LeuPheIsthrProThrGlnLysHsiIleGuaIalaLeuAArgGlnGlnIilaasnIalaIa | 280 |
| Db | | 939 | CTTTTTCACGCCTCCACTCAGAGACAATAGAACCCCTCCGGCAGAGATCATATCAACC | 998 |
| Oy | | 281 | ArgProGlnCySPProValGlyLeuAnthrIeulaIaPheProSerIlaAsnArgLyGlu | 300 |
| Db | | 999 | CGACCCCAGTCCCGCTGGGGCTTTAAACCCCTGGCTTCCACGATCAAACGGAAAGAA | 1058 |
| Oy | | 301 | ValValGluGluLysGlnProThrIalatyrtLeuSerCySGlyHisValHiGlyTyrrHis | 320 |
| Db | | 1059 | GTTGGGGAAGAGAAACAGCCCTGGGSCATACCTGACCTCGGCATGTCAAGGCTAACAC | 1118 |
| Oy | | 321 | AenTrpGlyHisiArgSerAspThrGluIalaenGluAArgGluCySPrometCyAsrThr | 340 |
| Db | | 1119 | AGCTGGGGCCATCCGAGGAGCAACGGAACCAACGAGAGAGTGTCCATGTGCAGACT | 1178 |
| Oy | | 341 | ValGlyProtyrValProLeuThrLeuGlyCySGlyValaglyPheTyrValAspAlaGly | 360 |
| Db | | 1179 | GTTGGCCCCCTACGTCCTCTCTGGGCTGTGTGAGCAGATTTTAAGTCGATGGGGA | 1238 |
| Oy | | 361 | ProProThrHisaIaPheThrProCySGlyHisiValCySeserGluYseSerAlaIasTy | 380 |
| Db | | 1239 | CCCCCACTCAGCGTTTCAACCCCTGGGGCAGCGTCTGTTCAGAGAAGTCTGGCAAATAC | 1298 |
| Oy | | 381 | TTPSerGlnIleProLeuProHisiGlyThrHisaIaPheHisiAlaAlaCySProPheCy | 400 |
| Db | | 1299 | TGGTGGCAGATCCCACTGCCACCGACGAAAGCAACGCGTTTCAACGCGCTGTCCGTTGC | 1358 |
| Oy | | 401 | AlaThrGlnLeuValGlyGluGlnAancCyBileLyLeuIlePheGlnGlyProIleAsp | 420 |
| Db | | 1359 | GCCAAGCAGCTGGTGTGTGAACAGAACTGCATCAATATGATTTTCCAAAGTCCAGTGCAC | 1418 |
| RESULT 5 | | | | |
| LOCUS | AK045673 | | 2564 bp | mRNA linear HTC 03-APR-2004 |
| DEFINITION | Mus musculus adult male corpora quadrigemina cDNA, RIKEN full-length enriched library, clone:B23020B21 product:pellino 1, full insert sequence. | | | |
| ACCESSION | AK045673 | | | |
| VERSION | AK045673.1 | | GI:26337532 | |
| KEYWORDS | HTC; CAP trapper. | | | |
| SOURCE | Mus musculus (house mouse) | | | |
| ORGANISM | Mus musculus | | | |
| REFERENCE | Eukaryaota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi; Muridae; Murinae; Mus. | | | |
| REFERENCE | 1 Carninci,P. and Hayashizaki,Y. | | | |
| AUTHORS | High-efficiency full-length cDNA cloning Mech. Enzymol. 303, 19-44 (1999) | | | |
| JOURNAL | PUBMED | | | |
| REFERENCE | 2 10349636 | | | |
| AUTHORS | Carninci,P., Shibata,Y., Hayatsu,N., Sugahara,Y., Shibata,K., Itoh,M., Komo,H., Okazaki,Y., Muramatsu,M. and Hayashizaki,Y. | | | |
| TITLE | Normalization and subtraction of cap-trapper-selected cDNAs to prepare full-length cDNA libraries for rapid discovery of new genes | | | |
| JOURNAL | PUBMED | | | |
| REFERENCE | 3 11042159 | | | |
| AUTHORS | Shibata,K., Itoh,M., Aizawa,K., Nagaoka,S., Sasaki,N., Carninci,P., ., | | | |

| | |
|-------------------|--|
| TITLE | THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE II TEAM AND THE PANTOM CONSORTIUM. |
| JOURNAL REFERENCE | FUNCTIONAL ANNOTATION OF A FULL-LENGTH MOUSE CDNA COLLECTION |
| AUTHORS | NATURE 409, 685-690 (2001) |
| TITLE | THE PANTOM CONSORTIUM AND THE RIKEN GENOME EXPLORATION RESEARCH GROUP PHASE I & II TEAM. |
| JOURNAL REFERENCE | ANALYSIS OF THE MOUSE TRANSCRIPTOME BASED ON FUNCTIONAL ANNOTATION OF 60,770 FULL-LENGTH CDNAS |
| AUTHORS | NATURE 420, 563-573 (2002) |
| TITLE | 6 (BASES 1 TO 2584) |
| JOURNAL REFERENCE | AADACH, T., AIZAWA, K., AKIMURA, T., ARAKAWA, T., BONO, H., CARNINCI, P., FUKUDA, S., FURUKO, M., HANAGAKI, T., HARA, A., HASHIZUME, M., HAYASHIZAKI, Y., HAYASE, N., HAYASE, N., HITOMOTO, K., HIROKAWA, T., HIROKAWA, T., HORI, F., IMOTO, K., ISHII, Y., ITOH, M., KAGAWA, I., KASUKAWA, T., KATO, H., KAWAI, J., KOJIMA, Y., KONDO, S., KONNO, H., KOURA, M., KOYA, S., KUNIHARA, C., MATSUYAMA, T., MIYAZAKI, A., MURATA, M., NAKAMURA, M., NISHI, K., NOMURA, K., NUMAZAKI, R., OHNO, M., OHSATO, N., OKAZAKI, Y., SATO, R., SATOH, K., SAKAI, C., SAKAI, K., SAKAZUME, N., SANO, H., SAKAKI, D., SHIBATA, K., SHINAGAWA, A., SHIRAKI, T., SOGABE, Y., TAGAMI, M., TAGAWA, A., TAKAHASHI, F., TAKAKU-AKASHITA, S., TAKEDA, Y., TANAKA, T., TOMARU, A., TOYA, T., YASUNISHI, A., YURAMATSU, M., AND HAYASHIZAKI, Y. |
| TITLE | DIRECT SUBMISSION |
| JOURNAL REFERENCE | SUBMITTED (16-JUL-2001) YOSHIIDE HAYASHIZAKI, THE INSTITUTE OF PHYSICAL AND CHEMICAL RESEARCH (RIKEN), LABORATORY FOR GENOME EXPLORATION RESEARCH GROUP, RIKEN GENOMIC SCIENCES CENTER (GSC), RIKEN YOKOHAMA INSTITUTE, 1-7-22 SUEHITO-CHO, TSURUMI-KU, YOKOHAMA, KANAGAWA 230-0045, JAPAN (E-MAIL: genome-res@gsc.riken.jp, URL: http://genome.gsc.riken.jp/, Tel: 81-45-503-9222, Fax: 81-45-503-9216) |
| COMMENT | CDNA LIBRARY WAS PREPARED AND SEQUENCED IN MOUSE GENOME ENCYCLOPEDIA PROJECT OF GENOME EXPLORATION RESEARCH GROUP IN RIKEN GENOMIC SCIENCES CENTER AND GENOME SCIENCE LABORATORY IN RIKEN. DIVISION OF EXPERIMENTAL ANIMAL RESEARCH IN RIKEN CONTRIBUTED TO PREPARE MOUSE TISSUES. |
| FEATURES | PLEASE VISIT OUR WEB SITE FOR FURTHER DETAILS. |
| SOURCE | URL: http://genome.gsc.riken.jp/ URL: http://fantom.gsc.riken.jp/. Location/Qualifiers |
| FEATURES | 1. .2584 /organism="Mus musculus" /mol_type="mRNA" /subset="CS7BL/6J" /db_xref="PANTOM_DB:B230220E21" /db_xref="taxon:10090" /clone="B230220E21" /sex="male" /tissue_type="corpora quadrigemina" /clone_id="RIKEN full-length enriched mouse cDNA library" /dev_stage="adult" /dev_1682 /note="unnamed protein product; pellino 1 (MD) [GI:1914495, GI NM_023324, evidence: BLASTN, 99%, match=2228] putative" /codon_start=1 /protein_id="BAC32452.1" /db_xref="GI:26337533" /translation="MPSDNDGNHPSKAPVYCGELIVYGVNGSLPNCGRPKSRPALE KRPKAVKSTYHIACTPDPAKAKISKDDHSISSTISRAQTVVETTHSDNTFQFQ GRTESSIDVVDITVFGSSNSDTSVOSTISIFACRIICERPPPTARIYAAGPDS |

Percent Similarity: 80.5%
 Best Local Similarity: 80.5%
 Query Match: 73.4%
 DB: 10
 Gaps: 0

US-10-041-030-4 (1-420) x AY409117 (1-1186)

27 TyranaglialaleuProasnglyAspArglyYargrglySeraRphalaLeuTyr 46
 2 TACAATGCTGCTTACCAATGAGATAGAGAGCGAGAAATAAGATTGGCTCTAC 61
 47 LysArgProlySalaanglyVallyeProSerThrValHsValIleSerThrProGln 66
 62 AAGGGCCCAAGGCAAAATGGTGCAAACCCGACATCTCCATGTGATATCCACGCCCCAG 121
 67 AlaserlySalaIleSerCyalyeGlyGlnHsSerIleSerThrLeuSerArgAsn 86
 122 GCGTCCAAAGGCTACAGCTGCMAAGGTCACACAGTATCTTACACTTGTCAAGGAAT 181
 87 GlnThrValValIgluTyrThrHsAspLySAspThrAspMetPheGlnValIgluArg 106
 182 CAGACTGTGGTGGTGGAGTACACATGATAGAGACACGATATGTTTCAGNNNNNNNN 241
 107 SerThrGluSerProIleAspPheValValThrAspThrIleSerGlySerGlnAsnThr 126
 242 NNN 301
 127 AspGluIaGlnIleThrGlnSerThrIleSerArgPheAlaCyAspGlyIleValCyAsp 146
 302 NNN 361
 147 ArgAsnGluProTyrThrAlaArgIlePheAlaIaGlyPheAspSerSerLyAsnIle 166
 362 NNN 421
 167 PheLeuGlyGlyValSalaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 186
 422 NNN 481
 187 ThrAsnGlyValLeuValMetHsProArgGlyGlyIleThrGluGluSerGlnProGly 206
 482 ACTAATGCGCTTCTGTGATGATCATCCAGAGGGGCTTCCACCGAGAGTCCACCGCGG 541
 207 ValTyrArgGluIleSerValCyGlyAspValTyrThrLeuAspGluThrArgSeraIa 226
 542 GTCTGGGGCGAGATCTGTCTGTGGAGATGTGTACCTTGGCAGAAACAGGTCCGCC 601
 227 GlnGlnArgGlyValLeuValGluSerGluThrAsnValLeuGlnAspGlySerLeuIle 246
 602 CAGCAACGAGAAAGCTGTGGAAGGAGACCAACGCTCCGACGAGCGGCTCCCTCAT 661
 247 AspLeuCyAspGlyAlaThrLeuLeuTyrArgThrAlaAspGlyLeuPheHsThrProThr 266
 662 GACCTGTGTGGGCGCTCTCTCTCTGAGAACACAGATGGGCTTTTCACTCCCACT 721
 267 GlnLysHsIleGluIa 286
 722 CAGAGACATAGAACGCTCTCCGCAAGATTAACGCCGCCGCTCCAGTGTCTGTG 781
 287 GlyLeuAsnThrLeuAlaPheProSerIleAsnArglySgluValIaIaIaIaIaIa 306
 782 GGGGCTCAACACCTGGGCTTCCACAGATCAACGAAAGAGGTGGTGGAGAGAAACAG 841
 307 ProTyrAlaTyrLeuSerCyGlyValIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 326
 842 CCCCTGGGCAATCTCACTGTGGCAGCTGCACGGGATCAACCACTGGGGCAATCGAGT 901
 327 AspThrGluIa 346
 902 GACACGAGGCGCAACGAGGAGGTGTCCNNNNNNNNNNNNNNNNNNNNNNNNNNNNNN 961
 347 LeuTyrLeuGlyCyGlyValIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 366

Db 962 CTCTGGCTGGCTGTGAGGAGGATTTTATAGACGACGACCGCCCACTNNNGCTTTC 1021
 Oy 367 ThrProCyGlyIa 386
 Db 1022 ACTCCCTGTGACACGTGTCTCGAGAGACTCTGCMAAATATCTGATCGATCCCTTG 1081
 Oy 387 ProHsGlyThrHsAlaPheHsAlaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIaIa 406
 Db 1082 CTTCAATGAATCATCATTCATTCACGCTGCTGCTTCTGTCTCAACAGCTGTGGG 1141
 Oy 407 GlnGlnAsnCyAlleIa 420
 Db 1142 GAGCAAAACTGCATCAATTAATTTTCCAAAGTCCAAATTGAC 1183

RESULT 7
 AK044418
 LOCUS
 DEFINITION
 AK044418 1879 bp mRNA linear HTC 03-APR-2004
 Mus musculus adult retina cDNA, RIKEN full-length enriched library,
 clone:A93001117 product:similar to PELLINO 2 (PELLINO (DROSOPHILA)
 HOMOLOG 2) (Homo sapiens), full insert sequence.
 AK044418.1 GI:26336455
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Mus musculus (house mouse)
 Mus musculus
 Mus musculus
 Mammalia; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Eukaryota; Euteleostomi; Euteleostomi; Euteleostomi; Euteleostomi;
 Sciurognathi; Muridae; Murinae; Mus.

REFERENCE
 AUTHORS
 TITLE
 JOURNAL
 PUBMED
 1 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Komuro, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 High-efficiency full-length cDNA cloning
 Meth. Enzymol. 303, 19-44 (1999)
 2 10349636
 3 Carninci, P., Shibata, Y., Hayatsu, N., Sugahara, Y., Shibata, K.,
 Itoh, M., Komuro, H., Okazaki, Y., Muramatsu, M., and Hayashizaki, Y.
 Normalization and subtraction of cap-trapper-selected cDNAs to
 prepare full-length cDNA libraries for rapid discovery of new genes
 Genome Res. 10 (10), 1617-1630 (2000)
 4 11042159
 5 Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carninci, P.,
 Komuro, H., Akiyama, J., Nishi, K., Kitsuai, T., Tashiro, H., Itoh, M.,
 Sumi, N., Ishii, Y., Nakamura, S., Hasegawa, T., Nishino, T., Harada, A.,
 Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K.,
 Fujisaki, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Matshid, M.,
 Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J.,
 Okazaki, Y., Muramatsu, M., Inoue, Y., Kira, A., and Hayashizaki, Y.
 RIKEN integrated sequence analysis (RISA) system-384-format
 sequencing pipeline with 384 multicapillary sequencer
 Genome Res. 10 (11), 1757-1771 (2000)
 6 11076861
 7 The RIKEN Genome Exploration Research Group Phase II Team and the
 FANTOM Consortium.
 Functional annotation of a full-length mouse cDNA collection
 Nature 409, 685-690 (2001)
 8 The FANTOM Consortium and the RIKEN Genome Exploration Research
 Group Phase I & II Team.
 Analysis of the mouse transcriptome based on functional annotation
 of 60,770 full-length cDNAs
 Nature 420, 563-573 (2002)
 9 (bases 1 to 1879)
 10 Adachi, J., Aizawa, K., Akimura, T., Arakawa, T., Bono, H., Carninci, P.,
 Fukuda, S., Furuno, M., Hanagaki, T., Hara, A., Hashizume, W.,
 Hayashida, K., Hayatsu, N., Hiramoto, K., Hiraoka, T., Hirozane, T.,
 Hori, F., Imotani, K., Ishii, Y., Itoh, M., Kagawa, I., Kasukawa, T.,
 Katoh, H., Kawai, J., Kojima, Y., Kondo, S., Komuro, H., Kouda, M.,
 Koya, S., Kurihara, C., Matsuyama, T., Miyazaki, A., Muraoka, M.,
 Nakamura, M., Nishi, K., Nomura, K., Numazaki, R., Ohno, M., Ohnaka, N.,
 Okazaki, Y., Saito, R., Saitoh, H., Sakai, C., Sakai, K., Sakazume, N.,
 Sano, H., Sasaki, D., Shibata, K., Shinagawa, A., Shiraki, T.,

FEATURES

Location/Qualifiers
 1. 2710
 /organism="Homo sapiens"
 /mol_type="mRNA"
 /db_xref="RZPD:DKFZp761p1024"
 /db_xref="taxon:9606"
 /clone="DKFZp761p1024"
 /tissue_type="amygdala"
 /clone_id="761 (synonym: hamy2). Vector pSPORT1; host DH10B; sites NotI + SalI"
 /dev_stage="adult"
 /note="pellino protein homolog 3"
 1. 2710
 /gene="DKFZp761p1024"
 112. 1521
 /gene="DKFZp761p1024"
 /product="hypothetical protein"
 /protein_id="CAD39057.1"
 /db_xref="GI:21740083"
 /db_xref="UniProt/Swiss-Prot:O8N2H9"
 /translation="MWLEGNPEVSPPTSDLOHRKNGSCVLSPEGDAOPGEPPIKGERLVLCCEGESETEAOREVETGPRAHSCYNGLASGDKRRSRSLASRSHANGVKPDVMMHISTPLYSKALSNRGOHSIYTSRSHSVVEYTHSDPMFOIGSTEMIDRVMDTSPGGGAABGSAOSTISRAACILCDRRPRAIYAAGDFSSMIFGEBAKMRTPDGLMDGLTTNGVLYVMHPRAGSESDAPGVREISVCANVTTRDSRAQKGLVENESNVLDGSLIDCGATLWMRTAGLRLAQLEAORQEAANAAPQCPVGLSTLAFSPARGRTAPDKOOPWYVACGHVGHGCRERGPORERDPLRLVGPVPLMLGOEAGLCIDPGRPSHAFAPCSHVSEKTAARYAQTPLPHGTATHAACPFCSAMLTGHCRLIFQGPLD"

ORIGIN

Alignment Scores:

Pred. No.: 6.66e-152 Length: 2710
 Score: 1620.00 Matches: 299
 Percent Similarity: 77.0% Conservative: 42
 Best Local Similarity: 67.5% Mismatches: 74
 Query Match: 70.7% Indels: 28
 DB: 4 Gaps: 4

US-10-041-030-4 (1-420) x HSM805479 (1-2710)

QY 3 SetProGlynglnuglnuGlyhScySalAProAsnlysgluproVallyrTyrglygluLeu 22
 Db 199 TCTCCCGGT-----GAAGATGCGCAGCCAGCGAGGACCCATCAATGATGGAATC 252
 QY 23 ValValleuGly----- 26
 Db 253 ATGCTCTGGGATGCTGTGAGAAAGAGGTGAGAAACCGAGGCTCGAGAGGGGAATG 312
 QY 27 -----TyrAsnGlyValaleuProAsnGlyAspArgGlyArg 38
 Db 313 ACTGGCCCAAGGCGACACAGCTGTACATGTGTGTGGCAAGTGGGGCAAGGGCGCG 372
 QY 39 ArgGlySerArgPheAlaLeuTyTyTyArgProLyAlaAsnGlyVallyPProSerThr 58
 Db 373 CGGCGAAGCGCGCTGCACTGAGCGCGGTGCGCAACCGGGGTGAAGCCAGACGTC 432
 QY 59 ValHisValIleSerThrProGlnAlaSerTyValIleSerCysLysglyGlnHisSer 78
 Db 433 ATGGACCAATCTCCAGCGCGCTGCTCCAGGCACTGAGTAACCGTGGTCAGCAACG 492
 QY 79 ILleSerTyThrIleuSerArgAsnGlnThrValValValGluTyThrHisAspLyAsp 98
 Db 493 ATCTCGATACACGCTCCCGAGCCACCTGCTCATATGAGTATACATGATAGGAC 552
 QY 99 ThrAspMetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAsp 118
 Db 553 ACAGACATGTTCCAGATTGGCCGCTCCACAGAGAACATGATGACTTCGTGGTAACAGAC 612
 QY 119 ThrIleSerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArg 138
 Db 613 ACGTCCCTGGAGAA---GGGGCTGCCAGAGGCCCTTCTGCGCCAGACACCATCTCCGC 669

QY 139 PheAlaCysArgIleValCysAspArgAsnGluProTyThrAlaArgIlePheAla 158
 Db 670 TATGCTGCGCGACCTCTGTGACCGCGCGACCCCTATATGCGCGCATATGCGCGCT 729
 QY 159 GlyPheAspSerSerLyAsnIlePheLeuGlyGluLyAlaAlaIleTyTrpLyAsnPro 178
 Db 730 GGTTCGATGCCCTTACCAACATCTTCTGGAGACGAGCGCCAAATGCGGACCCCA 789
 QY 179 AspGlyHisMetAspGlyLeuThrThrAsnGlyValleuValMetHisProArgGly 198
 Db 790 GATGCGCTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 849
 QY 199 PheThrGluGluSerGlnProGlyValIleTrpArgGluIleSerValCysGlyAspVal 218
 Db 850 TTCTCGAGAGCATCACCCCGCGGTGTGTGGCGGAAATCTCGTGTGGAAATGTGTAC 909
 QY 219 ThrLeuArgGluThrArgSerAlaGlnGlnArgGlyLyLeuValGluSerGluThrAsn 238
 Db 910 ACATTGGGAGACAGCCGCTCAGCCCGAGCGGAGAGCTGTGTGAAAACGATCCAC 969
 QY 239 ValLeuGlnAspGlySerLeuIleAspLeuCysGlyValaThrLeuLeuTrpArgThrAla 258
 Db 970 GTCTCGAGAGCGGCTCTCATGCACTGTGTGGGCCACACTGCTGTGTGGCACACCG 1029
 QY 259 AspGlyLeuPheHisThrProThrGlnLyGlnIleGluValAlaLeuArgGlnIleAsn 278
 Db 1030 GCGGCGCTGTGGCGGCTCCACATGAAAGCACTGAAGGCCAGCGGCGGAGGCAAT 1089
 QY 279 AlaAlaArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArg 298
 Db 1090 GCAAGCGGGGCCCGACAGTCCCGGTGAGCATCTGAGCTTCCCGACCGACGCCGT 1149
 QY 299 LyGluValValGluGluLys---GlnProTrpAlaTyLeuSerCysGlyHisValHis 317
 Db 1150 GCGCGCAGCGCGCGCGCAAAACAGCAGCCCTGGGTGTAGTGCGCGCGGACGTCAT 1209
 QY 318 GlyTyThrIleAsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMet 337
 Db 1210 GGCTACACAGCGCTGGGGCTGCCCGCGGAGCGGGCGCCCGACGAGAGCGCATCTCTTC 1269
 QY 338 CysArgThrValGlyProTyTrpValProLeuTrpLeuGlyCysGluAlaGlyPheTyVal 357
 Db 1270 TGCCGCTTGTGGGCGCTTATGTGCTCTATGCTTGGCCAGAGCGCGGCTTGGCTG 1329
 QY 358 AspAlaGlyProProThrHisAlaPheThrProCysGlyHisValCysSerGluTySer 377
 Db 1330 GACCTGGCGCGCTGACCATGCTTGTGACCTTGGGCGCACGCTGTGCTGTAGAAAGCT 1389
 QY 378 AlaLySerTyTrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaCys 397
 Db 1390 GCCCGTACTGGGCGACACACACATGCGCCACAGGCAACCATCTTCCATGCGGCTGC 1449
 QY 398 ProPheCysAlaIleThrGlnLeuValGlyGluGlnAsnCysGlyIleTyLeuIlePheGln 417
 Db 1450 CCTTTTGGCGGGCTGTGCTTACCGGCGAGCATGCGTGTGCTCATTTTCCAGGCGC 1509
 QY 418 ProIleAsp 420
 Db 1510 CCGCTGAT 1518

RESULT 12

AY408791

LOCUS

DEFINITION

genomic survey sequence.

ACCESSION

KEYWORDS

SOURCE

ORGANISM

Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;

Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;

[illegible][illegible]

| TITLE | JOURNAL | REFERENCE | AUTHORS |
|----------|---|---|--|
| 3 | Genome Res. 10 (10), 1617-1630 (2000) | 11042159 | Shibata, K., Itoh, M., Aizawa, K., Nagaoka, S., Sasaki, N., Carminci, P., Komoto, A., Akiyama, J., Nishi, K., Katsunai, T., Tashiro, H., Itoh, M., Suni, N., Ishii, Y., Nakamura, S., Hazama, M., Nishino, T., Harada, A., Yamamoto, R., Matsumoto, H., Sakaguchi, S., Ikegami, T., Kashiwagi, K., Fujikawa, S., Inoue, K., Togawa, Y., Izawa, M., Ohara, E., Watahiki, M., Yoneda, Y., Ishikawa, T., Ozawa, K., Tanaka, T., Matsura, S., Kawai, J., Okazaki, Y., Muramatsu, M., Inoue, Y., Kita, A. and Hayashizaki, Y. |
| TITLE | JOURNAL | REFERENCE | AUTHORS |
| 4 | Genome Res. 10 (11), 1757-1771 (2000) | 11076861 | RIKEN integrated sequence analysis (RISA) system--384-format sequencing pipeline with 384 multichiparray sequencer |
| TITLE | JOURNAL | REFERENCE | AUTHORS |
| 5 | Functional annotation of a full-length mouse cDNA collection | Nature 409, 685-690 (2001) | THE RIKEN Genome Exploration Research Group Phase II Team and the PANTOM Consortium. |
| TITLE | JOURNAL | REFERENCE | AUTHORS |
| 6 | Analysis of the mouse transcriptome based on functional annotation of 60,770 full-length cDNAs | Nature 420, 563-573 (2002) | Group Phase I & II Team. |
| TITLE | JOURNAL | REFERENCE | AUTHORS |
| 7 | Submitted (16-JUN-2001) | Yoshinori Hayashizaki, The Institute of Physical and Chemical Research (RIKEN), Laboratory for Genome Exploration Research Group, RIKEN Genomic Sciences Center (GSC), Kanagawa 230-0045, Japan [E-mail: genome-res@sc.riken.jp, Fax: 81-45-503-9216] | Direct Submission |
| COMMENT | cDNA library was prepared and sequenced in Mouse Genome Encyclopedia Project of Genome Exploration Research Group in Riken Genomic Sciences Center and Genome Science Laboratory in Riken. Division of Experimental Animal Research in Riken contributed to prepare mouse tissues. | | |
| FEATURES | <p>source</p> <p>1. 2474</p> <p>/organism="Mus musculus"</p> <p>/mol_type="mRNA"</p> <p>/strain="C57BL/6J"</p> <p>/db_xref="PANTOM_DB:9230114E15"</p> <p>/db_xref="taxon:10090"</p> <p>/clone="9230114E15"</p> <p>/sex="male"</p> <p>/tissue_type="epididymis"</p> <p>/clone_type="RIKEN full-length enriched mouse cDNA library"</p> <p>/dev_stage="adult"</p> <p><1. 854</p> <p>/note="unnamed protein product; pe11ino 2 (MGDI GJ:1891445, GB NW_033602, evidence: BLASTN, 99%, match=864)</p> <p>putative"</p> <p>/codon_start=3</p> <p>/protein_id="BAC28485.1"</p> | | |

```

/db_xref="GI:26329493"
/translation="RFAFCRIVCORBNPYTARIFAAGDSSKNIPIEKAKMKNPDGH
MDGTTTNGVLYVHPQGFTEESOPGWREISVCGDYITTLAETSAOQKGLVSENV
LQDSKILDLGGAITLMETADGLFHAIPQKIEALROBINARPOCPGNTLAPPSIN
RKVBERKQEPMAVYISGHVHSGHWRSTENNEBCEPCRWGPVPMICIEAF
YVAGPPTHAFTPCGHVCSERKSAKWSQPLPHGTHFAHAACPFCATQIVGEONCIDL
IYQSPVD"
polya_signal      2446..2451
                  /note="putative"
polya_site        2474
                  /note="putative"

ORIGIN
Alignment Scores:
Pred. No.:      3,32e-146      Length:      2474
Score:          1563.00        Matches:      279
Percent Similarity: 99.6%      Conservative: 3
Best Local Similarity: 98.6%    Mismatches: 1
Query Match:    68.3%          Indels:      0
Db:              4              Gaps:         0

US-10-041-030-4 (1-420) x AK033815 (1-2474)

Oy      138  ArpHeaIaCySaRgiLeValCySaPpArGaSngIuProTyThrIaLaRgiLePheAla 157
Db      3   AGGTTCCGATGCAAGATCGTGTGTGACAGAGACGACCAACGCAACACCAATTCGCG 62

Oy      158  AlAaGlyPheApsSerSerySaSnIlePheLeuGlyGluYsaIaIaLySTPLySaN 177
Db      63  GCAGATTCGATTCCTTCCAAAATATCTTCTTGAGAGAAGAACGACAAATGGAATAAC 122

Oy      178  ProAapGlyHISheLaPpGlyLeuThrThraSngIyAlleuValMetHISProARgLy 197
Db      123  CCTGATGACACAGATGATGACTCATACCAACCGTGTCTTGATGACACCCGCAAGGA 182

Oy      198  GlyPheThrgIuGluSergInProGlyValTTPArGgIuIleSerValCySgIyAPSPaJ 217
Db      183  GGGTTCCCGAGGAATCCAGCCTGAGGTCTGAGAGAGATCTGTGTGTGGAGATGTA 242

Oy      218  TyThrLeuARgIuThrArGserAlaGlnIaRgIyLySLeuValGluSergIuThr 237
Db      243  TACACCTTGGAGAGACCAAGTGGCCACAGAGGGGAAGCTGGTGAAGTGAAGACC 302

Oy      238  AsnValleuGlnAapGlySerLeuIleApsLeuCySgIyAlaThrLeuLeuTTPArGThr 257
Db      303  AACCTCCGCAAGCGGCTCCCTCATATGACCTGTGTGGGGCCACTCTCTCGAGAAACC 362

Oy      258  AlaAapGlyLeuPheHISThrProThrgIuYHISglIaIaIaIeuaRgIuGluIle 277
Db      363  GCAATATGCGCTTTTTCACGCTCTTAACCAAGACACATGAAGCCCTCCGGACAGAAATC 422

Oy      278  AsnAlaIaIaArGProGInCySProValGlyLeuAsnThrLeuAlaPheProSerIleAsn 297
Db      423  AATGACGCCCGAACCCCAAGTCCCGGTGGGCTTAAACCCCTGGCTTCCCAAGATCAAC 482

Oy      298  ArgLySgIuValValGluGluYsGInProTTPAlaTyLeuSeryCySgIyHISValHIS 317
Db      483  CGAAGAGAGTGGTGAAGAGAACCAAGCCCTGGGCATTAACCTGAGCTCGGCATGTGCAC 542

Oy      318  GlyTyThrHISaenTProGlyHISArgSerAapThrgIuAlaSnGluAaRgIuCySPromet 337
Db      543  GGGTACCAACAGCTGGGGCCATCGAGAGGACACGGAACCAACAGAGAGTGTCCATG 602

Oy      338  CyArGThrValAGlyProTyRValProLeuTTPLeuGlyCySgIuAlaGlyPheTyRVal 357
Db      603  TGCAGACATGTGGGCCCTTAAGTCCCTCTGGCTGGGCTGTAGAGCAAGATTTTAATGC 662

Oy      358  AspAlaGlyProProThrHISAlaPheThrProCySgIyHISValCySergIuYsSer 377
Db      663  GATCGGGAGACCCCAACTCAAGCTTTCAACCCCTCGCGGGACACGTGTTCAGAGAAGTCT 722

Oy      378  AlaIyeryTTPSergIuIleProLeuProHISGlyThrHISAlaPheHISAlaIaCyS 397

```


| | | | |
|---------------------------|--|---|-----------------------------|
| Db | 723 | GCAGAGTACTGATGTCAGATCCCACTGCGCCACGGAAGCAAGGATTTCACGCGCGCTGT | 782 |
| Oy | 398 | ProtheCySaIaTnThInLeuValGlYgluInIaNCyStIeYlVleuIlePheGInGly | 417 |
| Db | 783 | CCGTTTCTGCGCCACGACACTGATTGTGTAAACAGAACTGCATCAATGATTATTTTCCAAAGT | 842 |
| Oy | 418 | ProIleAaP | 420 |
| Db | 843 | CCACTGCAC | 851 |
| RESULT 14 | | | |
| LOCUS | BX388547 | 849 bp | mRNA linear EST 28-APR-2004 |
| DEFINITION | BX388547 Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED Homo sapiens | | |
| ACCESSION | BX388547 | | |
| VERSION | BX388547.1 | GI:30460950 | |
| SOURCE | EST. | | |
| ORGANISM | Homo sapiens (human) | | |
| REFERENCE | Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominiidae; Homo. | | |
| AUTHORS | Li, M.B., Gruber, C., Jeesee, J. and Polayes, D. | | |
| TITLE | Full-length cDNA libraries and normalization | | |
| JOURNAL | Unpublished (2001) | | |
| COMMENT | Contact: Genoscope Genoscope - Centre National de Sequencage 2 rue Gaspar Cremieux, CP 5706 - 91057 EVRY cedex - FRANCE Email: seqref@genoscope.cns.fr, Web : www.genoscope.cns.fr 1st strand cDNA was primed with a NotI-oligo (dT) primer. Five primers end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized. Library was constructed by Life Technologies, a division of Invitrogen. This sequence belongs to sequence cluster 7349.f | | |
| FEATURES | For more information about this cluster, see http://www.genoscope.cns.fr/cna?ts=CS0AT0082A06_T0713_1c=7349.f. | | |
| SOURCE | Location/Qualifiers | | |
| | 1..849 | | |
| | /organism="Homo sapiens" | | |
| | /mol_type="mRNA" | | |
| | /db_xref="taxon:9606" | | |
| | /clone="CS0DD006YF18" | | |
| | /cissue_type="NEUROBLASTOMA COT 50-NORMALIZED" | | |
| | /clone_lib="Homo sapiens NEUROBLASTOMA COT 50-NORMALIZED" | | |
| | /note="1st strand cDNA was primed with a NotI-oligo (dT) primer. Five prime end enriched, double-strand cDNA was digested with Not I and cloned into the Not I and EcoR V sites of the pCMVSPORT 6 vector. Library was normalized." | | |
| ORIGIN | | | |
| Alignment Scores: | | | |
| Pred. No.: | 1.8e-142 | Length: | 849 |
| Score: | 1519.00 | Matches: | 272 |
| Percent Similarity: | 98.6% | Conservative: | 1 |
| Best Local Similarity: | 98.2% | Mismatches: | 4 |
| Query Match: | 66.3% | Indels: | 0 |
| Db: | 5 | Gaps: | 0 |
| US-10-041-030-4 (1-420) x | BX388547 (1-849) | | |
| Oy | 140 | AlaCySaTgIleValCYaSPaRgASnGluPrOTyTThRaIaRgIlePheAlaAlaGly | 159 |
| Db | 1 | GCTTCGACATACGCTGTGCGACAGNATATACCTTCACAGCAGCGATATTTGCGCGCGGA | 60 |
| Oy | 160 | PheAaPSeSerIYtYASnIlePheLeuGlygluYsaIaAlaIaYtIryLYaSPaSPaP | 179 |
| Db | 61 | TTTGACTCTTCCAAAAACATATTTCTTGTGAGAAAAGCAGCAAAAGTGAAAAACCCCGAC | 120 |
| Oy | 180 | GIYhSWeTaaPGLYleuthTThRaNGIYValleuValMetHLePRAAGlyGlyPhe | 199 |

| Db | Accession | Gene | Species | Length (bp) | EST ID | EST Date |
|------------|--|---|---------|-------------|-----------------|----------|
| Db | 121 | GGCACAATGATAGGGCTCACTAATAAGGGTCTGTAATGATCAACAGAGGGGGCTTC | 180 | | | |
| Qy | 200 | ThgUgluSertInProglYvalITPhAggluIlEservAlYsgYlappValITrTh | 219 | | | |
| Db | 181 | ACCGAGGAGTCCACAGCCCGGGGCTCTGGCCGGAATCTCTGTCTGGAGATGTGTACCC | 240 | | | |
| Qy | 220 | LeuArGgluThraYrSerAlaInglInArYgGlyLeuValGluSertInTurAnVal | 239 | | | |
| Db | 241 | TTGGGAGAAACCAAGTCTGGCCCGACAGAGAAAGTGTGGAAATGAGACCAACGTC | 300 | | | |
| Qy | 240 | LeuGlnApsYlSerLeuIlleaPLeuCySglYalAthrLeuLeuTPArqThraAaSp | 259 | | | |
| Db | 301 | CTGCAGAGAGCGCTCCCTCATGTACCTGTGTGGGGCCACTCTCTTGAGAAACAGCAGAT | 360 | | | |
| Qy | 260 | GlyLeuPhehIThrProThrgInLYVhIleGluAlaLeuArGlnGluIlleaAnla | 279 | | | |
| Db | 361 | GGGCTTTTTCATCTCCAACTCAGAAAGCATGTAAAGCCCTCCGGACGAGATTAAAGCC | 420 | | | |
| Qy | 280 | AlaArGProGlnCyProValISglYLeuAnThrLeuAlaPheProSerIlleaAnArGlyS | 259 | | | |
| Db | 421 | GGCCGGCTCAAGTGTCTGTGGGGCTCAACACCCCTGCTCTCCACAGATCAACAGAA | 480 | | | |
| Qy | 300 | GlyValYalISgluGluYVgInProTPAlaTYrLeuSerCySglYhISvalHISglYTYr | 319 | | | |
| Db | 481 | GAGGTGTGAGAGAGAAAGACGCCCTGGCAATCTCAATGTTGTGGCACTGGCAAGGATAC | 540 | | | |
| Qy | 320 | HISaAnITPglYhISArGSeArPThrgluAlaAnGluuArGluCySPrometCyArG | 339 | | | |
| Db | 541 | CACAACTGGGGCCATCGAGATGACAGAGAGGCCAACAGAGGGAGATGTCCATGTGACGG | 600 | | | |
| Qy | 340 | ThrValISglYProTYrValPProLeuITPLeuGlyCySgluAlaISglYhetYrValaAPla | 359 | | | |
| Db | 601 | ACTGTGGGCCCTAATGTGCTCTCTGTGGCTTGGCTGTGAGGACAGAAATTTATGTAGACGA | 660 | | | |
| Qy | 360 | GlyProProThrHISAlaPheThrProCySglYhISvalCySertInLYVSeArAlaYs | 379 | | | |
| Db | 661 | GGACGCCCACTCAATGCTTCACTCCCTGTGGACAGATGTCTCGAGAGATGTGCAAAA | 720 | | | |
| Qy | 380 | TYrTPSerGlnIlleaProLeuProhISglYThrHISAlaPhehISAlaCySProphe | 399 | | | |
| Db | 721 | TACGTGCTCAGATCCCGTTGGCTCANTGAACTCANTGATTTACACCGTGTGCCCTTTC | 780 | | | |
| Qy | 400 | CyAlaThrGlnLeuValISgluGlnArCySIlleYsLeuIllePegin | 416 | | | |
| Db | 781 | TGTGTACACAGCTGGNTGGGAGACAAACTGCATCAATATTATTTTCAAA | 831 | | | |
| RESULT 15 | | | | | | |
| LOCUS | CX975821/c | | | | | |
| DEFINITION | CX975821 | 855 bp | mRNA | linear | EST 08-FEB-2005 | |
| ACCESSION | JGI_CAMP9337.fnd NIH XGC | | | | | |
| VERSION | IMAGE:7715625 5', mRNA sequence. | | | | | |
| KEYWORDS | CX975821 | | | | | |
| SOURCE | CX975821.1 | GI:58789343 | | | | |
| ORGANISM | EST. | | | | | |
| REFERENCE | Xenopus tropicalis (western clawed frog) | | | | | |
| AUTHORS | Xenopus tropicalis | | | | | |
| TITLE | Eukaryotic; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | | | |
| JOURNAL | Amphibia; Batrachia; Anura; Mesobatrachia; Pipridae; Pipidae; | | | | | |
| COMMENT | Xenopodinae; Xenopus; Silurana. | | | | | |
| | 1 (bases 1 to 855) | | | | | |
| | Richardson, P., Lucas, S., Rokhsar, D., Dettner, J.C., Ng, D.C., | | | | | |
| | Brokstein, P. and Lindquist, E.A. | | | | | |
| | DOB Joint Genome Institute Xenopus tropicalis EST project | | | | | |
| | Unpublished (2004) | | | | | |
| | Other_ESTRs: JGI_CAMP9337.rev | | | | | |
| | Contact: Lindquist, E.A., Richardson, P. | | | | | |
| | DOB Joint Genome Institute | | | | | |
| | 2800 Mitchell Drive, Walnut Creek, CA 94598, USA | | | | | |
| | Tel: 925 296 5600 | | | | | |
| | Fax: 925 296 5710 | | | | | |
| | Email: cdna@jgi-psf.org | | | | | |
| | Tissue Procurement: Robert M. Grainger | | | | | |
| | cDNA Library Preparation: Bruce Blumberg Laboratory, University of | | | | | |

California Irvine
DNA Sequencing: DOE Joint Genome Institute: <http://www.jgi.doe.gov>
Clone Distribution: I.M.A.G.E. Consortium/LNL:
<http://image.lnl.gov>
Naming Conventions: EST name is generated by the concatenation of the JGI Clone id and the direction of sequencing. The suffix '.fw' indicates a forward sequencing read of the insert. It does not necessarily reflect the orientation of the insert.
plate: CMAP 0097 row: b column: 7
High quality sequence stop: 807.

FEATURES

Source

```

/organism="Xenopus tropicalis"
/mol_type="mRNA"
/strain="N6 (Nigerian 6th generation inbred)"
/db_xref="taxon:8364"
/clone="IMAGE:7715625"
/tissue_type="Intestine"
/dev_stage="Adult"
/lab_host="Electromax DH10B T1 Phage Resistant cells"
/clone_id="NH XGC triop1"
/notes="Vector: PCS107; Site 1: EcoRI; Site 2: XhoI; The library was prepared from 5 ug of poly A+ RNA by oligo-dT priming"
(5'-GAGAGAGAGAGAGAGAGACTGTCGAGTTTATTTTATTTT-3')
and
Stratascript reverse transcriptase. After ligation of EcoRI adapters (5'-ATTCCGACAGAG-3') followed by kinsing adapters and by XhoI digestion, the cDNA was size selected by chromatography on Sepharose CL-2B columns and fractions containing cDNAs larger than 1000 bp were ligated into EcoRI/XhoI-digested PCS107. Reference for library construction: Current Genomics 4, 635-644. Library constructed by Michelle Tabb and Bruce Blumberg (Dept of Developmental and Cell Biology, University of California, Irvine)."
```

ORIGIN

| Alignment Scores: | | | |
|---|-----------|---------------|-----|
| Pred. No.: | 4,84e-140 | Length: | 855 |
| Score: | 1495.00 | Matches: | 265 |
| Percent Similarity: | 96.8% | Conservative: | 10 |
| Best Local Similarity: | 93.3% | Mismatches: | 9 |
| Query Match: | 65.3% | Indels: | 0 |
| DB: | 8 | Gaps: | 0 |
| US-10-041-030-4 (1-420) x CX975821 (1-855) | | | |
| QY 134 SerThrIleSerArgPheAlaCysArgIleValCysAspArgAsnGluProTyrThrAla | | | 153 |
| 853 AGTACATCTCTGATTTTCGTCGAGATAGCTGTGACAGGAACCTCCGTACACGCT | | | 794 |
| QY 154 ArgIlePheAlaAlaIlePheAspSerIleValIlePheLeuGluIleValAla | | | 173 |
| 793 AGATATNTGCTGCAGAGATTGACCTCCCAAAAATATATTTCTTGTGGAGAAAGCGGC | | | 734 |
| QY 174 LysTrpLysAsnProAspGlyHisMetAspGlyLeuThrThrAsnGlyValIleValMet | | | 193 |
| 733 AATGTGAAGATCTGTATGTCTATATGATGATGACATACCAATGCGTCTTGTATG | | | 674 |
| QY 194 HisProArgGlyGlyPheThrGluGluSerGlnProGlyValTTPArgGluIleSerVal | | | 213 |
| 673 CATCCCAAGAGGAGGGTTTACTGACAGATCCAGCGCTGAGTATGGCGGAAATTTCAATC | | | 614 |
| QY 214 CysGlyAspAlaTyrThrLeuArgGluTyrArgSerAlaGlnGlnArgGlyLysLeuVal | | | 233 |
| 613 TGTGGAGATATATATCATTTAGAGAAACCAATCTGCTCAGCANAGAGCAATCTGTG | | | 554 |
| QY 234 GluSerGluThrAsnValLeuGlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeu | | | 253 |
| 553 GAAATGGAACCAATTTCTGCGAGATGGTCTCTTGTGACTTATATGGTGCACACCTT | | | 494 |
| QY 254 LeuTPArgThrAlaAspGlyLeuPheHisThrProThrGlnLysHisIleGluAlaLeu | | | 273 |

| | | | |
|----|-----|---|-----|
| Db | 493 | CTTGGCGAAACAGCAGATGTTGGTTCATACACCGACTCAAAAACATGGAAGCTTTG | 43 |
| Qy | 274 | ArgGInGluIleAsnAlaIaIaArgProGlnCysProValaGlyLeuAsnThrLeuAlaPhe | 293 |
| Db | 433 | AGACTGAAATTAAGTCCAGAGCCTCAGTGCCTGTTGATTAAACATTTAGCTTTT | 374 |
| Qy | 294 | ProSerIleAsnArgGlySerIleValaGlnGluIlyGlnProTrrAlaIyTrLeuSerCys | 313 |
| Db | 373 | CCAGATATCAACCGCAAGATAGATGCTGTAAAGAGAAACGCTTGGGCTTACCTCAGCTGT | 314 |
| Qy | 314 | GlyHisValaHisGlyTyrrHisAsnThrProIyHisArgSerAsnThrGlnAlaAsnGluArg | 333 |
| Db | 313 | GGTCATGTCACAGGCTATCATACCTGGGGTCAATGCATGACACAGAAATCATGAAGCG | 25 |
| Qy | 334 | GluCysArgProMetCysAsnThrValGlyProTyrValaProLeuThrPleuGlyCysGlyAla | 353 |
| Db | 253 | GAATGCCATGATGTGAGAAACAGTTGCTGTCCTGATGAGCTGTATGGCTTGGTTGGTGAAGCA | 194 |
| Qy | 354 | GlyPheTyValaIaAspAlaGlyProProThrHisAlaPheThrProCysGlyHisValaCys | 373 |
| Db | 193 | GGATTTTACGTGAAGTCCAGACCTTCTACTACAGCTTTCAAACCTTGTGGACAGCTAGC | 134 |
| Qy | 374 | SerGluIySerAlaIyTyrrTrpSerGlnIleProIeuProHisGlyThrHisAlaPhe | 393 |
| Db | 133 | TCTGAAAAATCTCCAAATATGGTCTAGATTCACATGCCACATGGTATCAACAGCGCTTT | 74 |
| Qy | 394 | HisAlaAlaCysArgProPheCysAlaIaThrGlnLeuValGlyGlnIleAsnCysIleGlyLeu | 413 |
| Db | 73 | CATGCTGCTTGTCTTTCTTGCAGAAATCAATATGACCTGGAGAAAGAAAGCTGTGTCAAACTT | 14 |
| Qy | 414 | IlePheGlnGly 417 | |
| Db | 13 | ATTTCCTCAGGCG 2 | |

Search completed: March 2, 2006, 09:03:06
Job time : 5371 Secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - nucleic search, using frame_plus.p2n model

Run on: March 2, 2006, 04:45:41 ; Search time 809 Seconds

(without alignments)
3460.038 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MESPGEHAPKPEPVKVG.....ATQVGEQNCIKILPQPID 420

Scoring table: BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delext 7.0

Searched: 4996997 seqs, 3332346308 residues

Total number of hits satisfying chosen parameters: 9993994

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=xlh
-O=/abs/ABSWEB.spool/US10041030/runat_01032006_134403_21809/app.query.fasta-1
-DB=N Geneseg -OFFT=fastap -SUPPLX=p2n.rng -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blonum62 -TRANS=human40.cdi -LIST=45
-DOCNALGN=200 -THR_SCORE=pct -THR_MAX=100 -THR_MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NOR=ext -HEADSIZE=500 -MINLEN=0 -MAXLEN=200000000 -HOST=abs808
-USER=US10041030 @CGN 1.1 727 @runat_01032006_134403_21809 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSPLOCK=100 -LONGLOG -DEV TIMEOUT=120
-MAIN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

N_Geneseg_21: *
1: geneseg1980s: *
2: geneseg1990s: *
3: geneseg2000s: *
4: geneseg2001s: *
5: geneseg2001bs: *
6: geneseg2002as: *
7: geneseg2002bs: *
8: geneseg2003as: *
9: geneseg2003bs: *
10: geneseg2003cs: *
11: geneseg2003ds: *
12: geneseg2004as: *
13: geneseg2004bs: *
14: geneseg2005s: *

Pred. No. is the number of results predicted by chance to have a
score greater than or equal to the score of the result being printed,
and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----|---------------------|
| 1 | 2290 | 100.0 | 1565 | 6 | AAL46323 Human M33 |
| 2 | 2290 | 100.0 | 1823 | 10 | ADG30209 Human nov |
| 3 | 2290 | 100.0 | 4563 | 6 | ABQ78319 Nucleotid |
| 4 | 2290 | 100.0 | 5579 | 8 | ACC42349 Human MAMP |

| | | | | | |
|----|--------|------|-------|----|---------------------|
| 5 | 2283 | 99.7 | 1263 | 6 | ABL58449 Human pel |
| 6 | 2283 | 99.7 | 1263 | 12 | ADP48670 Human pel |
| 7 | 2175.5 | 95.0 | 1260 | 6 | ABL58448 Human pe |
| 8 | 2175.5 | 95.0 | 1260 | 12 | ADP48668 Mouse pel |
| 9 | 2111.5 | 92.2 | 1717 | 8 | ACC42348 Mouse MAP |
| 10 | 1917 | 83.7 | 1257 | 6 | ABL58447 Human pel |
| 11 | 1917 | 83.7 | 1257 | 12 | ADP48666 Human pel |
| 12 | 1917 | 83.7 | 1304 | 8 | ABX05095 Human nov |
| 13 | 1917 | 83.7 | 3222 | 6 | AAL46320 Human M30 |
| 14 | 1917 | 83.7 | 3526 | 6 | AAL46319 Human M30 |
| 15 | 1917 | 83.7 | 7136 | 12 | ADP83102 Human pel |
| 16 | 1917 | 83.7 | 7136 | 13 | ADP56419 Human PRO |
| 17 | 1917 | 83.7 | 7136 | 13 | ADP24928 PRO polyP |
| 18 | 1917 | 83.7 | 7136 | 14 | ADY17389 DNA encod |
| 19 | 1917 | 83.7 | 7136 | 14 | ADY20613 DNA encod |
| 20 | 1916 | 83.7 | 2736 | 6 | AAL46316 Rat M30 c |
| 21 | 1912 | 83.5 | 1257 | 6 | ABL58446 Murine pe |
| 22 | 1912 | 83.5 | 1257 | 12 | ADP48664 Mouse pel |
| 23 | 1822 | 79.6 | 3846 | 6 | AAL46317 Human M30 |
| 24 | 1642 | 71.7 | 1338 | 6 | ABL58452 Human pel |
| 25 | 1642 | 71.7 | 1338 | 12 | ADP48674 Human pel |
| 26 | 1642 | 71.7 | 2683 | 10 | ADP71199 Human int |
| 27 | 1640 | 71.6 | 2589 | 6 | AAL46322 Human M31 |
| 28 | 1632 | 71.3 | 2588 | 10 | ADG30836 Human nov |
| 29 | 1632 | 71.3 | 1936 | 6 | AAL46321 Human M3 |
| 30 | 1618 | 70.7 | 3254 | 6 | AAL46318 Human M30 |
| 31 | 1605 | 70.1 | 3332 | 3 | AAC66453 Human sec |
| 32 | 1574 | 68.7 | 2866 | 4 | AAH16312 Human CDN |
| 33 | 1574 | 68.7 | 2866 | 4 | AAH77808 Nucleotid |
| 34 | 1482.5 | 64.7 | 2508 | 10 | ADH63804 Human CDN |
| 35 | 1370 | 59.8 | 2183 | 4 | AAK94855 Human ful |
| 36 | 1370 | 59.8 | 2183 | 12 | ADJ31993 Pull leng |
| 37 | 1302 | 56.9 | 2966 | 4 | ABL09073 Drosophill |
| 38 | 1291 | 56.4 | 16092 | 4 | AAK73420 Human imm |
| 39 | 1172 | 51.2 | 803 | 4 | AAK73420 Human imm |
| 40 | 1102 | 48.1 | 592 | 10 | ADG32115 Human nov |
| 41 | 1075 | 46.9 | 224 | 5 | AAK73420 Human imm |
| 42 | 1064 | 46.5 | 567 | 6 | AAL46314 Human M33 |
| 43 | 1050 | 45.9 | 570 | 6 | AAL46324 Human M32 |
| 44 | 1022 | 44.6 | 541 | 12 | ACH89787 Human gen |
| 45 | 996 | 43.5 | 31882 | 4 | ABL09072 Drosophill |

ALIGNMENTS

RESULT 1
ID AAL46323 standard; cDNA; 1565 BP.
AC AAL46323;
DT 19-JUL-2002 (first entry)
XX
DE Human M33 coding sequence SEQ ID NO: 15.

XX Neurodegenerative disease; M30; M31; M32; M33; stroke;
XX fragile X syndrome; Huntington's disease; Parkinson's disease;
XX Alzheimer's disease; multiple sclerosis; ovarian cancer;
XX neurodegeneration; immune disorder; autoimmune disease; allergy;
XX infection; leukaemia; inflammation; neuroprotection; cerebroprotective;
XX immunosuppressive; cytostatic; nootropic; antiparkinsonian; antiallergic;
XX virucide; antiinflammatory; gene; ss.

OS Homo sapiens.

PN NO200221138-A2.

PD 14-MAR-2002.

PF 07-SRP-2001; 2001NO-BP010366.

PR 07-SRP-2000; 2000US-00657479.

XX 10-APR-2003.
 PD 24-SEP-2002; 2002MO-US030474.
 XX 24-SEP-2001; 2001US-0324631P.
 XX (HYSB-) HYSB INC.
 PA Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Drmanac RT;
 XX WPI: 2003-371981/35.
 DR P-PSDB; ADC31180.
 XX New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 PS Claim 1; SEQ ID NO 291; 1185bp; English.
 XX The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the invention
 CC (ADC31394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human cDNA sequence of the invention. Note: The sequence data for
 CC this patent did not form part of the printed specification, but was
 CC obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.
 XX
 SQ Sequence 1823 BP; 425 A; 483 C; 530 G; 385 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 9.69e-200 Length: 1823
 Score: 2290.00 Matches: 420
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 10 Gaps: 0
 US-10-041-030-4 (1-420) x ADC30209 (1-1823)
 QY 1 MetPheSerProGlyGlnGluGlnHisCysAlaProAsnGlyGluProValIleTyrGly 20
 Db 265 ATGTTTCCCTGGCGAGGAACTGCGCCCAATAGAGGAGCCAGTGAATACGGG 324
 QY 21 GluLeuValIleGlnGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgLys 40

Db 325 GAGCTGTGGTGTCTGGGTATCAATGTGCTTACCAATGAGATGAGAGCGAGGAAA 384
 QY 41 SerArgPheAlaLeuTyrIleAspArgProIleAlaAsnGlyValIleProSerThrValHis 60
 Db 385 AGTAGATTTGGCCCTTCAAGCCGCCCAAGGCGMAATGTGTCTCAAAACCCAGCACCGTCCAT 444
 QY 61 ValIleSerThrProGlnAlaSerGlyAlaIleSerCysLysGlyGlnHisSerIleSer 80
 Db 445 GTATATTCACGGCCCAAGGATCAAGGCTATCAGCTGCAGAAAGTCAACAGATATATCC 504
 QY 81 TyrThrLeuSerArgAsnGlnThrValIleValIleGlyTyrThrHisAspLysAspThrAsp 100
 Db 505 TACACTTTCAGAGATCAGACTGTGTGTGTGTGATACACATATAGATATAGGAT 564
 QY 101 MetPheGlnValIleArgSerThrGluSerProIleAspPheValIleAspThrIle 120
 Db 565 ATGTTTCAGGTGGGAGATCAACAGAAAGCCATATGACTTCGTTGTACAGACAGATT 624
 QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 625 TCTGGAGCGCAACACGAGCGAAGCCAGATTCACAGAGCACATATCCAGGTTCCG 684
 QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160
 Db 685 TCGAGATCTGTGTCCACAGATGAACTTACACAGCAGATATTCGCCCGGATTT 744
 QY 161 AspSerSerLysAsnIlePheLeuGlyGlyValAlaIleTyrLysAsnProAspGly 180
 Db 745 GACTCTTCCAAAACATATTTCTTGGAGAAAGGACGAAAGTGAAACCCGACGGC 804
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
 Db 805 CACATGATGTGGCTCACTACTATATGCGCTGTGTATGATCATCAGAGGGGCTTCACC 864
 QY 201 GluGluSerGlnProGlyValITrPAArgLuuIleSerValCysGlyAspValIYrThrLeu 220
 Db 865 GAGAGATCCAGCCCGGGGTCTGGCCGAGATCTGTCTGTGAGATGTGTACACTTG 924
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
 Db 925 CGAGAAACACAGTGTGGCCCGACGAAACAGGAAAGCTGTGGAAGTGAGACCAAGTCTTG 984
 QY 241 GluAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTrpArgThrAlaAspGly 260
 Db 985 CAGAGCGGCTCCCTCATTTGACTGTGTGGGGCCACTCTCTCTGAGAACAGCAGATGGG 1044
 QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAlaIle 280
 Db 1045 CTTTTTCATATCTCAACTCAAGACATAGAAAGCCCTCCGAGAGATTAAGCCCGCC 1104
 QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
 Db 1105 CGGCTCAAGTCTCTGTGGGGCTCAACACCTGTGCTTCCCGACATCAACAGAAAGAG 1164
 QY 301 ValValGlnGluLysGlnProTrpAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
 Db 1165 GTGTGAGAGGAGAAAGCGCTGTGGCATATCTCAGTTGTGGCCACCGTGCAGGTTACAC 1224
 QY 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
 Db 1225 AACCTGGGGCATGTGAGTGCACGAGGCGCAACAGAGGAGGTGTCCATGTCCAGACT 1284
 QY 341 ValGlyProTyrValProLeuTrpLeuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
 Db 1285 GTGGGCCCTATGTGCTCTGTGCTGTGTGTGAGAGGAGATTTATGTAGAGCGAGGA 1344
 QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlyLysSerAlaLysTyr 380
 Db 1345 CCGCCAACTATATCTTCACTCCCTGTGACACGTGTGTCTGGAGAAAGTCTGCAAAATAC 1404
 QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaIleCysProPheCys 400

QY 361 ProProThHisAlaPheThrProCysGluValCysSerGluYSerAlaLysIYr 380
 DB 1581 CCGCAACTCATCTTCACTCCCTGTGACAGCTGTCTGGAGAGCTCGCAAAATAC 1640
 QY 381 TPSSerGlnIleProLeuProHISglYThHisAlaPheHISAlaAlaCysProPheCys 400
 DB 1641 TGGTCTCAGATCCCGTTGGCTCATGGAACATGACATTCACGCTGCTTCCCTTCTGT 1700
 QY 401 AlaThGlnLeuValGluGlnAsnCysIleLysLeuIlePheGlnIYProIleAsp 420
 DB 1701 GCTACACAGCTGGTGGGAGCAAACTGCATCAAAATTAATTTCCAGTCCCAATTGAC 1760
 RESULT 4
 ACC42349
 ID ACC42349 standard, cDNA; 5579 BP.
 AC ACC42349;
 XX
 DT 22-MAY-2003 (first entry)
 DE Human MAP kinase cascade activator #59 cDNA.
 XX
 XX Human; Elki phosphorylation; Elki phosphorylation kinase; virucide;
 KM antiinflammatory; immunomodulator; cyostatic; anti-allergic; anti-HIV;
 KM antineumatic; antiarthritis; antidiabetic; antistematic; gene therapy;
 KM inflammation; autoimmune disease; viral disease; cancer; diabetes;
 KM rheumatoid arthritis; asthma; allergic rhinitis; AIDS; viral hepatitis;
 KM IGA nephritis; gene; ss.
 XX
 OS Homo sapiens.
 XX
 PN MO2003008589-A1.
 XX
 PD 30-JAN-2003.
 XX
 PF 15-JUL-2002; 2002MO-JP007174.
 XX
 PR 18-JUL-2001; 2001JP-00218204.
 PR 31-AUG-2001; 2001JP-00263450.
 PR 21-JAN-2002; 2002JP-00012176.
 XX
 PA (ASAH) ASAH KASEI KOGYO KK.
 XX
 PI Matsuzaki O, Matsuda A, Nagano Y, Suzuki N;
 DR WPI; 2003-229582/22.
 DR P-PSDB; ABR41083.
 XX
 PT Elki phosphorylation-associated gene and its encoded protein with MAP
 PT kinase cascade effect, applicable in diagnosis of and developing drugs
 PT for e.g. inflammation, autoimmune diseases, viral diseases and cancer.
 XX
 PS Claim 4; Page 631-640; 762pp; Japanese.
 XX
 CC The invention relates to a novel purified protein having Elki
 CC phosphorylation activity and/or an activity of activating Elki
 CC phosphorilation kinase. A protein of the invention has antiinflammatory,
 CC immunomodulator, virucide, cyostatic, antiallergic, antineumatic,
 CC antineumatic, antidiabetic, antistematic, and anti-HIV activity. The
 CC polynucleotides may have a use in gene therapy. The gene and its encoded
 CC protein are applicable in diagnosis of and developing drugs for e.g.
 CC inflammation, autoimmune diseases, viral diseases and cancer such as
 CC rheumatoid arthritis, diabetes, asthma, allergic rhinitis, AIDS, viral
 CC hepatitis and IGA nephritis. The present sequence is used in the
 CC exemplification of the invention
 XX
 SQ Sequence 5579 BP; 1535 A; 1111 C; 1214 G; 1719 T; 0 U; 0 Other;

Alignment Scores:
 Pred. No.: 4,46e-199
 Score: 2290.00
 Percent Similarity: 100.0%
 Best Local Similarity: 100.0%

Length: 5579
 Matches: 420
 Conservative: 0
 Mismatches: 0

Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0
 US-10-041-030-4 (1-420) x ACC42349 (1-5579)
 QY 1 MetPheSerProGluYngInGluGlnHisCysAlaProAsnLysGluProValLysIYrGly 20
 DB 177 ATGTTTCCCTCGGCGAGAGAACACTGCGCCCAATAGAGAGCAATGAAATACGGG 236
 QY 21 GluLeuValIleuGluYTyraGngIYalaLeuProAsnGlyAspArgGlyArgArgLys 40
 DB 237 GAGCTGTGGTGTGGGTGATCAATGTGTCTTAACCAATGAGATAGAGAGCGAGAA 296
 QY 41 SerArgPheAlaLeuTyryLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
 DB 297 AGTAGATTGGCTCTTACAGCGCCCAAGGCAAAATGGTGTCAAAACCGACCGTCCAT 356
 QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGluYngInHisSerIleSer 80
 DB 357 GTGATATCCAGCCCGAGGATCCAGGCTTACGCTGCAAGGTCAACACAGTATATCC 416
 QY 81 TyThrLeuSerArgAsnGlnThrValValIleGluTyThrHisAspLysAspThrAsp 100
 DB 417 TACACTTTGTCAAGAAATCAGACTGTGTGTGGAGTACACACATGATAGATACGAT 476
 QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValIleThraPThrIle 120
 DB 477 ATGTTTCAGGTGGGAGATCAACAGAAAGCCCTATGACTTCTGTTCACACAGATT 536
 QY 121 SerGlySerGlnAsnThrAspGlnIleGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 537 TCTGGAGCGCAAGAACGAGACAGAGCCAGATCACAGAGCACATATCCAGGTTCGCC 596
 QY 141 CysArgIleValCysAspArgAsnGluProTyThrAlaArgIlePheAlaIleGlyPhe 160
 DB 597 TGCAGATGCTGTGGCAGACAGAAATGAACCTTACAGACAGATATTCGCCGCGGATT 656
 QY 161 AspSerSerLysAsnIlePheLeuGlyGluValAlaAlaLysTyrLysAsnProAspGly 180
 DB 657 GACTCTTCCAAAMAACTATTTCTTGAGAAAGGCGCAAGGAGGAAAAACCCCGACGCC 716
 QY 181 HisMetAspGluLeuThrThrAsnGlyValIleValMetHisProArgGlyGlyPheThr 200
 DB 717 CACATGAAATGGGTCACTACTAATGCGCTCTGTATGATCATCAGAGGGGCTTCACC 776
 QY 201 GluGluSerGlnProGluYValTyrArgGluIleSerValCysGluYAspValTyThrLeu 220
 DB 777 GAGGATCCCAAGCCCGGGGTCTGGCCGAGATCTCTGTGAGATGTGTACACTTG 836
 QY 221 ArgGluThrArgSerLagInGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
 DB 837 CGAGAAACAGAGTCGAGCCAGCAACAGAAAGCTGTGAAAGTGAAGACCAACGTCCTG 896
 QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuThrArgThrAlaAspGly 260
 DB 897 CAGAGAGGCTCCCTCTTAACCTGTGTGGGCGACCTCTCTCGAGAACAGAGAGTGG 956
 QY 261 LeuPheHisThrProThrgInLysHisIleGluValLeuArgGlnIleAlaAla 280
 DB 957 CTTTTCATATCTCAACTCAGAGACATAGAGACCTCCGAGAGAGATTAACGCGCC 1016
 QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
 DB 1017 CGGCTCAGGTGCTGTGGGGCTCAACACCTGCGCTTCCCGCATCAACAGGAAGAG 1076
 QY 301 ValValIleGluGluLysGlnProThrPalatyTyLeuSerCysGlyHisAlaHisGlyTyHis 320
 DB 1077 GTGGTGAGAGAGAGAGAGAGCCCTGGGCAATCTCAGTTGTGGCCACGTGACGGGTACAC 1136
 QY 321 AsnTrGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
 DB 1137 AACCTGGGCGCATGAGTACACAGAGCGCAACAGAGAGGAGATGTCTCATGTCCAGACT 1196

| | | | |
|----------|---|--|------|
| Qy | 341 | ValGlyProCysTrpValProLeuThrLeuGlyCysGlyValAspGlyPheTrpValAspAlaGly | 360 |
| Db | 1197 | GTGGGCCCCCTAATGAGCCCTCTCTGGCTTGAGCTGTGGAGGAAATTTTATGTAGACGCAGGA | 1256 |
| Qy | 361 | ProProThrHisAlaPheThrProCysGlyHisValCysSerGlyLysSerAlaLeuTyr | 380 |
| Db | 1257 | CCGCCAATCATGATGCTTTCACTCCCTGTGGACACCGTGCTCGAGAAAGTCTGCAAAATAC | 1316 |
| Qy | 381 | TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys | 400 |
| Db | 1317 | TGGTCTCAGATCCCGTCCTTCATGAACTCATGCTTTCACGCGCTGCTTCCTTCGT | 1376 |
| Qy | 401 | AlaThrGlnLeuValAlaGlyGluAsnGlyIleLeuLeuIlePheGlnGlyProIleAsp | 420 |
| Db | 1377 | GCTACACAGCTGGTGTGGGAGCAAAACTGCATCAATTAATTTTCCAAAGTCCAAATTGAC | 1436 |
| RESULT 5 | | | |
| ID | ABL58449 | ABL58449 standard; DNA; 1263 BP. | |
| AC | ABL58449; | | |
| XX | 30-JUL-2002 | (first entry) | |
| XX | Human pellino-2 | polypeptide coding sequence. | |
| XX | Nuclear factor-kappaB; NF-kappaB; p38; pellino; interleukin-1; virucide; | | |
| KW | antibacterial; fungicide; protozoicide; antiasthmatic; antihemetic; | | |
| KW | antiarthritic; antiinflammatory; antiatherosclerotic; neuroprotective; | | |
| KW | nocotropic; antidiacer; human; pellino-2; gene; ds. | | |
| XX | Homo sapiens. | | |
| OS | Key | Location/Qualifiers | |
| XX | Key | 1..1263 | |
| XX | CDS | /*tag= a | |
| FT | /product= "pellino-2" | | |
| FT | WO200183739-A2. | | |
| PN | 08-NOV-2001. | | |
| XX | 27-APR-2001; 2001WO-US013676. | | |
| XX | 28-APR-2000; 2000US-0200198P. | | |
| XX | (IMMUNEX CORP. | | |
| PA | Bird TA, Coeman DJ; | | |
| PI | WPI; 2002-066532/09. | | |
| XX | P-PSDB; ABB07922. | | |
| DR | New Pellino polypeptides for identifying compounds that alter polypeptide | | |
| PT | activity, treating pathogenic infection or inhibiting apoptosis, are | | |
| PT | capable of stimulating nuclear factor-kappaB- or p38-dependent | | |
| PT | transcription. | | |
| XX | Claim 3; Page 61; 70pp; English. | | |
| XX | The invention provides polypeptides capable of stimulating nuclear factor | | |
| CC | (NF)-kappaB-dependent transcription or p38-dependent transcription, | | |
| CC | referred as Pellino polypeptides. The pellino polypeptides are useful for | | |
| CC | identifying modulators that alter the pellino polypeptide and pellino | | |
| CC | dominant-negative activity. They are also useful for identifying | | |
| CC | compounds that inhibit the binding activity of the polypeptides and to | | |
| CC | study cell-signal transduction. They are useful for preventing or | | |
| CC | treating infection by a pathogen such as virus, bacterial, fungi, algae | | |
| CC | or protozoa, or inhibiting apoptosis. Dominant-negative pellino | | |
| CC | polypeptides are useful for treating inflammatory conditions such as | | |
| CC | asthma, rheumatoid arthritis, inflammatory bowel disease, Crohn's | | |
| CC | disease, ulcerative colitis, atherosclerosis and Alzheimer's disease, and | | |
| CC | also for inhibiting mitogen activated protein (MAP) kinase-activated | | |

CC pathways; Pellino polypeptides and polyunucleotides are useful to identify
CC small molecule inhibitors of protein association or function of Pellino,
CC and other molecules involved in interleukin (IL)-1 signalling. The present
CC sequence represents a human pellino-2 polypeptide coding sequence
XX
SQ Sequence 1263 BP, 317 A, 336 C, 351 G, 259 T, 0 U, 0 Other;

Alignment Scores:

| | |
|------------------------|-----------------|
| Pred. No.: | 2,576=199 |
| Score: | 2283.00 |
| Percent Similarity: | 99.8% |
| Best local Similarity: | 99.8% |
| Query Match: | 99.7% |
| DB: | 6 |
| | Gaps: 0 |
| | Matches: 419 |
| | Conservative: 0 |
| | Mismatches: 1 |
| | Indels: 0 |
| | Gaps: 0 |

US-10-041-030-4 (1-420) x ABL58449 (1-1263)

1 MetpHeSerProGlyGlnGluGluHisCysAlaProAsnLysGluProValLysTyrGly 20

Db 1 ATGTTTCCCTGGCCAGGAGACACTGGCCCCCAATAGAGACCAGTGAATACGGG 60

21 GluLeuValValIleuGlyTyrAsnGlyValAlaLeuProAsnGlyAspArgGlyArgArgLys 40

Db 61 GAGCTGGTGGTCTCGGGTACATGTGCTTTACCCAATGAGATGAGGACCGAGGAAA 127

41 SerArgPheAlaLeuTyrIysArgProIysAlaAsnGlyValIysProSerThrValHis 60

Db 121 AGTAGTTTGCCCTCTACAGCGGCCCAAGCAATGTGTCAAACCCAGCACCGTCCAT 180

61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80

Db 181 GTGATATCCACGCCCCAGGCAATCCAGGCTATCAGTGCAGCAAGGTCAACACAGTATATCC 241

81 TyrThrLeuSerArgAsnGlnThrValValValGlnTyrThrHisAspLysAspThrAsp 10

Db 241 TACACTTGTCAAGGATCAGCTGTGCTGTGAGTACACACATGATTAAGGATACGGAT 30

101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 122

Db 301 ATGTTTCAGGTGGCAGATCA CAGAAAGCCCTATCGACTTCGTTGTCA CAGACAGATT 36

QY 121 SerGlySerGlnAsnThrAspCysAlaGlnIleThrGlnSerThrIleSerArgPheAla 14

Db 361 TCTGAGCCAGAACGAGCAGGCCAGTCAACAGAGCACCATTCCAGTTCCGCC 422

141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIaGlyPhe 16

Db 421 TGCAGATCGTGTGCGACAGGATGAACTTACACAGCAGGATATGCGCCGCGATTT 48

161 AspSerSerLysAsnIlePheLeuGlycylulysAlaAlaLysTrpLysAsnProAspGly 18

Db 481 GACTCTTCAAAAACATATTTCTTGAGTAAAGCAGCAAGTGAAAAACCCGACGC 54

181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyIlePheThr 20

Db 541 CACATGGATGGGCTCACTACTAATGGCGTCTGGTGATGCATCCAGAGGGGCTTACC 60

QY 201 GIUGLuserGIInProGIyValITrpargGIuIeserValCyBGIyaspValITyrThrlLeu 22

Db 601 GAGAGTCCACGCCGGGGTCTGGCGGAGATCTCTGTCTGTGGAGATGTATACACTTG 666

221 ArgGIuThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 24

Db 661 CGAGAACCAAGTCG66CCACGACGAGGAAAGCTGGTGAAGAAGTGAGACCAACGTCCTG 72

241 GlnaspGlySerLeuIleaspLeuGlyAlaThrLeuLeuTrpArgThrIleaspGly 26

Db 721 CAGGACGGCTCCCTCATTTGACTGTGTGGGGCCACTCTCCTCTGGAGAACAGACAGATGGG 78

261 LeuPheHisThrProThrglnLysHisIleGluAlaLeuArgGlnGluIleAsnAla 28

Db 781 CTTTTCATCTCCA CTGAGA GCACTAGAGCCCTCCGGCAGAGATTAA CGCGCC 84

QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLySGlu 300
 Db 841 CGGCTCAGTGTCTCTGGGGCTCAACCTGCTTCCCGACATCAACAGAAAGAG 900
 QY 301 ValValGlnGluLysGlnProTTPAlaTyrLeuSerCysGlyValIHisGlyTyrHis 320
 Db 901 GTGGTGGAGAGAGACAGCCCTGGGCAATCTCATGTTGGCCACGTGCACGGGTATCAC 960
 QY 321 AsnTTPGlyHisArgSerArgThrGluAlaAsnGluArgLysProMetCysArgThr 340
 Db 961 AACTGGGGCCATCGAGTGAACGGAGGCCAACAGAGGAGTGTCCCATGTGACACT 1020
 QY 341 ValGlyProTyrValProLeuTTPLeuGlyCysGlnValGlyPheTyrValAspAlaGly 360
 Db 1021 GTGGGCCCCATATGTGCTCTCTGCTGGCTGTGAGCAGAGATTATTTATGACCGCAGA 1080
 QY 361 ProProThrHisAlaPheThrProCysGlyValIHisValCysSerGluLysSerAlaLysTyr 380
 Db 1081 CCGGCAACTCATGCTTTCACCTCCCTGTGACAGCTGTGCTCGAGAGAGTGTGCAAAATAC 1140
 QY 381 TTPSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
 Db 1141 TGGTCTCAGATCCCGTGTGCTCATGGAACCTCATGCTCATGCTGCTGCTTCTGT 1200
 QY 401 AlaThrGlnLeuValGlyGlnGluAsnValIleLysLeuIlePheGlnGlyProIleAsp 420
 Db 1201 GCTACACAGCTGTTGGGGAGCAAACTGATCAATTAATTTCCAAAGTCCCAATGAC 1260
 RESULT 6
 ADP48670 ID ADP48670 standard; cDNA; 1263 BP.
 AC ADP48670;
 XX 09-SEP-2004 (first entry)
 DT
 DE Human Pellino-2 encoding cDNA SEQ ID NO:7.
 XX
 XX nuclear factor kappa B dependent transcription inhibitor;
 KW NF-kB-dependent transcription inhibitor;
 KW p38-dependent transcription inhibitor; Pellino-1;
 KW Interleukin 1 receptor-associated kinase 4;
 KW IL-1 receptor-associated kinase 4; IRAK-4; antiarteriosclerotic;
 KW antiarthritic; antiasthmatic; antiinflammatory; antirheumatic; antiulcer;
 KW Gastrointestinal; neuroprotective; nootropic; IL-1 antagonist;
 KW IRAK protein kinase family inhibitor;
 KW NF kappa B inducing kinase inhibitor; asthma; rheumatoid arthritis;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW atherosclerosis; Alzheimer's disease; human; Pellino-2; chromosome 14;
 KW gene; ss.
 XX
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 1..1263
 FT /*tag= a
 FT /product= "Pellino-2"
 PN MO2004053092-A2.
 XX
 PD 24-JUN-2004.
 XX
 PF 09-DEC-2003; 2003WO-US039188.
 XX
 PR 11-DEC-2002; 2002US-00317250.
 XX
 PA (IMMUNEX CORP.
 PA (CLEV-) CLEVELAND CLINIC FOUND.
 PI Bird TA, Cosman DJ, Li X;
 XX
 DR WPI: 2004-480927/45.
 DR P-PDB: ADP48671.

XX
 PT Identifying inhibitors of IL-1 signaling, useful for treating e.g.,
 PT asthma, and rheumatoid arthritis, comprises assaying the association of a
 PT Pellino-1 polypeptide with IRAK-4 in the presence of a test compound.
 PS Claim 9; SEQ ID NO 7; 81dp; English.
 XX
 CC The present invention describes a method for identifying compounds (C)
 CC that inhibit nuclear factor kappa B (NF-kB)-dependent transcription or
 CC p38-dependent transcription. The method comprises: (1) mixing a test
 CC compound with a Pellino-1 polypeptide; (b) assaying the association of
 CC the Pellino-1 polypeptide with IL-1 receptor-associated kinase 4 (IRAK-4)
 CC in the presence of the test compound; and (c) determining whether the
 CC test compound inhibits the association of the Pellino-1 polypeptide with
 CC a binding partner. Also described: (1) an inhibitory nucleic acid that
 CC binds to a nucleic acid encoding an amino acid sequence as described
 CC above, where the presence of the inhibitory nucleic acid within a cell
 CC inhibits the association of Pellino-1 with IRAK-4; and (2) an inhibitory
 CC polypeptide comprising an antibody fragment that binds to a polypeptide
 CC comprising an amino acid sequence as described above, where the presence
 CC of the inhibitory polypeptide within a cell inhibits the association of
 CC Pellino-1 with IRAK-4. (C) have antiarteriosclerotic, antiarthritic,
 CC antiasthmatic, antiinflammatory, antirheumatic, antiulcer,
 CC gastrointestinal, neuroprotective and nootropic activities, and can be
 CC used as an IL-1 antagonist, IRAK protein kinase family inhibitor, and NF
 CC kappa B inducing kinase inhibitor. The methods and compositions of the
 CC present invention are useful for the prevention and/or treatment of
 CC diseases or conditions associated with aberrant expression or activity of
 CC the Pellino-1 polypeptide, such as asthma, rheumatoid arthritis,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC atherosclerosis and Alzheimer's disease. The present sequence encodes
 CC human Pellino-2, which is used in the exemplification of the present
 CC invention. The human Pellino-2 gene is located on chromosome 14, more
 CC specifically to 14q24.3.
 XX
 SQ Sequence 1263 BP; 317 A; 336 C; 351 G; 259 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 2,57e-199 Length: 1263
 Score: 2283.00 Matches: 419
 Percent Similarity: 99.8% Conservative: 0
 Best Local Similarity: 99.8% Mismatches: 1
 Query Match: 99.7% Indels: 0
 DB: 12 Gaps: 0
 US-10-041-030-4 (1-420) x ADP48670 (1-1263)
 QY 1 MetPheSerProGlnGlnGluGlnHisCysAlaProAsnLysGluProValLysTyrGly 20
 Db 1 ATGTTTCCCTCGCGGAGAGAACCTGGGCCCAATAGAGAGCAAGTGAATACGGG 60
 QY 21 GluLeuValValLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgLysArgTglys 40
 Db 61 GAGCTGTGTGTCTCGGATCAATGATGCTTACCAATGAGATAGAGAGCGAGAAA 120
 QY 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
 Db 121 AGTAGATTGGCCCTCTCAAGCGCGCCCAAGGCAAAATGCTCAAAACCGACACGCTTCAT 180
 QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlnGlyLysIleSer 80
 Db 181 GTGATATCCACGCCCAAGGATCCAAAGCTATAGCTGCAAGGTCAACACAGTATATCC 240
 QY 81 TyrThrLeuSerArgAsnGlnThrValValValGlyTyrThrHisAspLysAspThrAsp 100
 Db 241 TACACTTGTCAAGATCAAGATCACTGTGTGTGATGATCAACATATAGATAGGAT 300
 QY 101 MetPheGlnValGlyArgSerThrGlySerProIleAspPheValValThrAspThrIle 120
 Db 301 ATGTTTCAGGTGGGAGATCAACAGAAAGCCCATGATGACTGCTGTCAACAGACGATT 360
 QY 121 SerGlySerGlnAsnThrAspGluValGlnIleThrGlnSerThrIleSerArgPheAla 140

```

Db      361 TCTGGCAGCCAGAACGAGCAAGCCAGATCA CACAGAGCACATATTCAGGTTGCC 420
Qy      141 CAAAGTllevel CAAAPATgAengluProTyThralaaglllephelaalaglyPhe 160
Db      421 TCGAGGATCTGTGCGACGAGATGAACCTTACACAGCAGATATTCGCCCGGAGTTT 480
Qy      161 AASPserlysaenllepheleuglyglulysalalalytTplysaenProaspGly 180
Db      481 GACTCTCCAAAAAATATTTCTTGAGTAAAGGCACAAAGTGAAAAAAGCCCGACGGC 540
Qy      181 HlmechaspelyleuthrthraenglyValleuValmechlsProaaglygllyPheThr 200
Db      541 CACATGATGGGCTCACTACTAAATGGCGCTGCTGTATGATCATCCACAGAGGGGCTTCACC 600
Qy      201 GlugluserglInProgllyValtTPATgglulleserValCysegllyasppVallyrThleu 220
Db      601 GAGGAGTCCAGCCCGGAGGTGCGGAGATCTCTGTCTGTGAGATGTGTACACCTTG 660
Qy      221 ArggluthrarseseralaglnglnaraglylylsleuValglusergluthrAsnValleu 240
Db      661 CGAGAAACAGGTGCGGCCAGCAACGAGAAAGCTGTGGAAGTGAGACCAAGTCTCG 720
Qy      241 GluaspGlyserleuileaspLeuCysegllyAlathrleuLeuthrPATgthralaaspGly 260
Db      721 CAGGACGGCTCCCTCATTTGACCTGTGTGGGGCCACTCTCTCTGGAGAACAGCAGATGGG 780
Qy      261 LeupheithrProthrglnlyshlsllaglnalaleuarglnglnulleaenlaala 280
Db      781 CTTTTTATCTCCACTCAGAACACATAGAACCCCTCCGGCAGGAGATTAACGCCGCC 840
Qy      281 ArgProglInCyseProValgllyleuAsnThrleuAlapheProserlleaenArglyglu 300
Db      841 CGGCTCAGTGTCTGTGGGGCTCAACACCTGGCTTCCCGACATCAACAGAAAGAG 900
Qy      301 ValValgluglulysglInProthralaTytleuserCysegllyshlsllaglyrThls 320
Db      901 GTGGTGAAGAGAGAGAGCCCTGGGCAATCTCAGTTGTGGCCACGTGCACGGGTACAC 960
Qy      321 AsmtTgplYhlsArgSerAspThrglnAlaAsnlnuargluCyseProMetCyseArgThr 340
Db      961 AACTGGGGCCATGCGAGTGAACGAGAGCCCAACGAGAGAGTCTCCATGTGCAAGACT 1020
Qy      341 ValgllyProTyThralProleuthrleuglyCysegllyalaglyPheTyThralaaspAlgly 360
Db      1021 GTGGGGCCCTATGAGCTCTCTGGCTTGGCTGTGAGGACGAGATTATTAATGAGCGAGGA 1080
Qy      361 ProProthrhlsalalaphethrProCysegllyshlsValCyseSerglulysSerAlalyrTy 380
Db      1081 CCGCCAACTCATGTCTTCACTCCCTGTGGAACCTGTGCTCGGAGAGTCTGCAAAATAC 1140
Qy      381 TrpSerGlnlleProleuProHlsaglYThrhlsalalaphethlsalalaaCyseProPheCys 400
Db      1141 TGGCTCAGATCCCGTGGCTTCAGAGAACTCAGCATTTCAAGCTGTGCCCTTCTGTGT 1200
Qy      401 AlathrglnleuValgllygluglnAsnCyselleysleuilepheglnnglyProlleasp 420
Db      1201 GCTAACAGCTGGTGGGAGCAAACTGCATCAATTAATTTTCCAAGTCCCATTTGAC 1260

```

RESULT 7

ABLS8448 standard; DNA; 1260 BP.

ABLS8448;

30-JUL-2002 (first entry)

Murine pellino-2 coding sequence.

Nuclear factor-kappaB; NF-kappaB; p38; pellino; interleukin-1; virucide;
 antiinfectious; fungicide; protozoacide; antiasthmatic; antirheumatic;
 antitubercular; antiinflammatory; antitubercular; antiparasitic; neuroprotective;
 nootropic; anticancer; mouse; pellino-2; gene; ds.

XX

OS Mus musculus.

XX Key Location/Qualifiers

XX Key CDS 1..1257

XX FT /tag= a

XX FT /product= "pellino-2"

XX PN MO200183739-A2.

XX PD 08-NOV-2001.

XX XX 27-APR-2001; 2001MO-US013676.

XX XX 28-APR-2000; 2000US-0200198P.

XX XX (IMMUNEX CORP.

XX XX Bird TA, Cosman DJ;

XX XX WPI; 2002-06532/09.

XX XX P-PSDB; ABB07921.

XX XX Example; Page 58-59; 70pp; English.

XX CC The invention provides polypeptides capable of stimulating nuclear factor

XX CC (NF)-kappaB-dependent transcription or p38-dependent transcription,

XX CC referred as Pellino polypeptides. The pellino polypeptides are useful for

XX CC identifying molecules that alter the pellino polypeptides and pellino

XX CC dominant-negative activity. They are also useful for identifying

XX CC compounds that inhibit the binding activity of the polypeptides and to

XX CC study cell-signal transduction. They are useful for preventing or

XX CC treating infection by a pathogen such as virus, bacterial, fungi, algae

XX CC or protozoa, or inhibiting apoptosis. Dominant-negative pellino

XX CC polypeptides are useful for treating inflammatory conditions such as

XX CC asthma, rheumatoid arthritis, inflammatory bowel disease, Crohn's

XX CC disease, ulcerative colitis, atherosclerosis and Alzheimer's disease, and

XX CC also for inhibiting mitogen activated protein (MAP) kinase-activated

XX CC pathways. Pellino polypeptides and polymucleotides are useful to identify

XX CC small molecule inhibitors of protein association or function of Pellino,

XX CC and other molecules involved in interleukin (IL)-1 signaling. The present

XX CC sequence represents a murine pellino-2 polypeptide coding sequence

XX SQ Sequence 1260 BP; 309 A; 357 C; 359 G; 235 T; 0 U; 0 Other;

XX Alignment Scores:

XX Pred. No.: 1,836-189 Length: 1260

XX Score: 2175.50 Matches: 399

XX Percent Similarity: 97.4% Conservative: 10

XX Best Local Similarity: 95.0% Mismatches: 10

XX Query Match: 95.0% Gaps: 1

XX DB: 6 Gaps: 1

US-10-041-030-4 (1-420) x ABLS8448 (1-1260)

Qy 1 MetPheSerProgllynglnglnuHlsCyalaProAsnlygluProVallyrTyrgly 20

Db 1 ATGTTTCCCGGGCGAGAGAAACCAAGGCGCCCAACAGAGAGCGGTGAATACGGG 60

Qy 21 GluleuValValleuglyTyraenglyAlaleuProAsnGlyAspArglyValaaglyls 40

Db 61 GAGCTGTGTCTCTGTGAGTAAATGCTTAACTTAATGTGTAACGAGGCGAGAGGAAA 120

Qy 41 SerArgPheAlaLeuTyrlsArgProlyAlaAsnGlyVallyrProserThValHls 60

Db 121 AGCAGATTGCTCTTATTAAGCGACTTACGAGGTGTCAAAACCGACAAATCCAC 180

Qy 61 ValIleSerThrProglInlaserlyAlaIleSerCyalegllyglInIleSer 80

XXXXXX

Db 181 ATGGTCTCCACACACAGCGCTCCAGGCGCATGCTCCAGAGCATCACAGCATATCG 240
 QY TyrThrLeuSerArgAsnGlnThrValValGluTyrThrHisAspIleAspThrAsp 100
 Db 241 TACACGTTGTCACGAGCGACGAGCGTAGTGATGATACACACGATTAAGACACCGAC 300
 QY MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
 Db 301 ATCTTTACAGTGGGAGAGTCAACAGAAAGCCCATTTACCTTGGTGGTCAACAGACGCGTT 360
 QY SerGlySerGlnAsnThrAspGluValGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 361 TCCGCGCGGTGACAGACGAAAGAT---GCCCATGATCAACAGAGCAACATCTCTAGGTTGCA 417
 QY CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaIleGlyPhe 160
 Db 418 TGCAGAGATCGTGTGTACAGAAACGAGCCATATACAGCAGCATATTCGCGGACGATTC 477
 QY AspSerSerIleAsnIlePheLeuGlyGluValAlaIleValIleTyrPheAsnProAspGly 180
 Db 478 GATTTCTCCAAAATATCTTTCTTGAGAGAAAGACCAAAATGAAAAACCTTGATGGA 537
 QY HisMetAspGlyLeuThrThrAsnGlyValIleValMetHisProArgGlyGlyPheThr 200
 Db 538 CATATGATGATGATCACTACCAATGGTGTCTGTATGATGACCCGCAAGAGAGGCTTACC 597
 QY GluGluSerGlnProGlyValIleTyrArgGluIleSerValCysGlyAspValIleThrLeu 220
 Db 598 GAGGAATCCCGAGCTGAGATGAGAGATCTCTGTGTGGGAGTGTGTACACCTTG 657
 QY ArgGlnThrArgSerIleAsnGlnIleArgGlyLeuLeuValGluSerGluThrAsnValLeu 240
 Db 658 CGAGAGCCAGGTCTGCGCCACAGAGGGGAAAGCTGTGAAAGTGAACCAACGTCCTG 717
 QY GlnAspGlySerLeuIleAspLeuCysGlyValAlaThrLeuLeuThrArgThrAlaAspGly 260
 Db 718 CAGACGGGCTCCCTCATTTGACTGTGTGGGCGCACTCTCTCGAGAAACCGAGATGGC 777
 QY LeuPheHisThrProThrGlnIleHisIleGluIleLeuArgGlnGluIleAsnAlaAla 280
 Db 778 CTTTTCACGCTCTCTACTCAGAAACATAGAAAGCCCTCCGCGACGAGATCAATGCGACC 837
 QY ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgGlyGlu 300
 Db 838 CGAGCCCAAGTGCCTCGGGGCTTTAACACCTGCGCTTCCCGACATCAACCGAAGGAA 897
 QY ValValGluGluValProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
 Db 898 GTGGTGGAAAGAGAGAGCCCTCGGCACTACCTGAGCTCGGCAATGTCACGGCTTACC 957
 QY AsnTyrGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
 Db 958 AGCTGGGGCCATCGAGCGAGCGAGGAGCAAGAGGAGGAGTGTCCATGTGAGAGACT 1017
 QY ValGlyProTyrValProLeuTyrPheGlyCysGluAlaGlyPheTyrValAlaAspAlaGly 360
 Db 1018 GTGGGGCCCCACCTCTCTCTGTGGTGGTGTGAGGCGAGATTTATGTCGATGCGGGA 1077
 QY ProProThrHisAlaPheThrProCysGlyHisValCysSerGluValSerAlaValTyr 380
 Db 1078 CCCCACATCAACGCTTACCCCTCGGCGAGAGCTCTGTTCAGAAAAGTCTGCCAAGTAC 1137
 QY TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
 Db 1138 TGGTCCAGATCCCATCGCCCAACGGAACGCAACGCTTATGCGCTGCTGCGCTTGC 1197
 QY AlaThrGlnLeuValGlyGluGlnAsnCysIleLeuLeuIlePheGlnGlyProIleAsp 420
 Db 1198 GCCACGACGCTGTGTGTGAACGAACTGCATCAAAATGATTTCCAAAGGTCCAGTGGAC 1257

RESULT 8
 ADP48668
 ID ADP48668 standard; cDNA; 1260 bp.

XX AC ADP48668;
 XX DT 09-SEP-2004 (first entry)
 DE Mouse Pellino-2 encoding cDNA SEQ ID NO:5.
 XX KW nuclear factor kappa B dependent transcription inhibitor;
 KW NF-kB-dependent transcription inhibitor; Pellino-1;
 KW p38-dependent transcription inhibitor; Pellino-1;
 KW interleukin 1 receptor-associated kinase 4;
 KW IL-1 receptor-associated kinase 4; IRAK-4; antiarteriosclerotic;
 KW antiarthritic; antiasthmatic; antiinflammatory; antirheumatic; antitumor;
 KW gastrointestinal; neuroprotective; nootropic; IL-1 antagonist;
 KW IRAK protein kinase family inhibitor;
 KW NF kappa B inducing kinase inhibitor; asthma; rheumatoid arthritis;
 KW inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 KW atherosclerosis; Alzheimer's disease; mouse; Pellino-2; gene; ss.
 XX OS Mus musculus.
 XX FH Key
 XX FT CDS
 XX FT 1. 1260
 XX FT /tag= a
 XX FT /product= "Pellino-2"
 XX MO2004053092-A2.
 XX PD 24-JUN-2004.
 XX PF 09-DEC-2003; 2003MO-US039188.
 XX PR 11-DEC-2002; 2002US-00317250.
 XX PA (IMMUNEX CORP.
 XX PA (CLEV-) CLEVELAND CLINIC FOUND.
 XX PI Bird TA, Cosman DJ, Li X;
 XX WPI; 2004-480927/45.
 XX DR P-PSDB; ADP48669.
 XX PT Identifying inhibitors of IL-1 signaling, useful for treating e.g.,
 PT acthina, and rheumatoid arthritis, comprises assaying the association of a
 PT Pellino-1 polypeptide with IRAK-4 in the presence of a test compound.
 PS Example 1; SEQ ID NO 5; 81pp; English.
 XX The present invention describes a method for identifying compounds (c)
 CC that inhibit nuclear factor kappa B (NF-kB)-dependent transcription or
 CC p38-dependent transcription. The method comprises: (1) mixing a test
 CC compound with a Pellino-1 polypeptide; (b) assaying the association of
 CC the Pellino-1 polypeptide with IL-1 receptor-associated kinase 4 (IRAK-4)
 CC in the presence of the test compound; and (c) determining whether the
 CC test compound inhibits the association of the Pellino-1 polypeptide with
 CC a binding partner. Also described: (1) an inhibitory nucleic acid that
 CC binds to a nucleic acid encoding an amino acid sequence as described
 CC above, where the presence of the inhibitory nucleic acid within a cell
 CC inhibits the association of Pellino-1 with IRAK-4; and (2) an inhibitory
 CC polypeptide comprising an antibody fragment that binds to a polypeptide
 CC comprising an amino acid sequence as described above, where the presence
 CC of the inhibitory polypeptide within a cell inhibits the association of
 CC Pellino-1 with IRAK-4. (c) have antiarteriosclerotic, antiarthritic,
 CC antiasthmatic, antiinflammatory, antirheumatic, antitumor,
 CC gastrointestinal, neuroprotective and nootropic activities, and can be
 CC used as an IL-1 antagonist, IRAK protein kinase family inhibitor, and NF
 CC kappa B inducing kinase inhibitor. The methods and compositions of the
 CC present invention are useful for the prevention and/or treatment of
 CC diseases or conditions associated with aberrant expression or activity of
 CC the Pellino-1 polypeptide, such as asthma, rheumatoid arthritis,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC atherosclerosis and Alzheimer's disease. The present sequence encodes
 CC mouse Pellino-2, which is used in the exemplification of the present

CC protein are applicable in diagnosis of and developing drugs for e.g.
 CC inflammations, autoimmune diseases, viral diseases and cancer such as
 CC rheumatoid arthritis, diabetes, asthma, allergic rhinitis, AIDS, viral
 CC hepatitis and IGA nephritis. The present sequence is used in the
 CC exemplification of the invention

Sequence 1717 BP; 395 A; 483 C; 501 G; 324 T; 0 U; 14 Other;

Alignment Scores:

| Pred. No.: | 2.06e-183 | Length: | 1717 |
|------------------------|-----------|---------------|------|
| Score: | 211.50 | Matches: | 388 |
| Percent Similarity: | 95.2% | Conservative: | 12 |
| Best Local Similarity: | 92.4% | Mismatches: | 19 |
| Query Match: | 92.2% | Indels: | 1 |
| DB: | 8 | Gaps: | 1 |

US-10-041-030-4 (1-420) x ACC42348 (1-1717)

```

QY 1 MetPheSerProGlyGlnGluGluHisCysAlaProAsnLysGluProValLysTyrGly 20
DB 218 ATGTTTCCCGGCGCAGAGAACCCAGCGCCCAAGAGCGCGTGAATACAGG 277
QY 21 GluLeuValAlaLeuGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgGly 40
DB 278 GACCTGGTGGTCTCGGGGTACANTGTCCTTACCTAATGTCACAGGGCAGAGAGAAA 337
QY 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
DB 338 AGCAGATTATCCCTCTAAGCCGACCTACGCCAGTGTCTCAAAACCAACACATCCAC 397
QY 61 ValIleSerThrProGlnAlaSerLysValIleSerCysLysGlnHisSerIleSer 80
DB 398 ATGCTCTCCACACCACAGCGGCTCCAGGCCATCACTCCAGCAGATCAAGCATATCG 457
QY 81 TyrThrLeuSerArgAsnGlnThrValValAlaGluTyrThrHisAspLysAspThrAsp 100
DB 458 TACACGTTGTCACGAGCCAGCCAGCTGTGTGTGAAGTACACACGATTAAGACACAGAC 517
QY 101 MetPheGlnValGlyArgSerThrGlnSerProIleAspPheValAlaThrAspThrIle 120
DB 518 ATGTTTCAGGTGGGACAGTCAACAGAAAGCCCATTCATTCGTGTCACAGACACGGTT 577
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 578 TCCGGCGGTGAGAACGAAAT--GCCAGATCAACAGAGCCACATCTCTAGGTTGCA 634
QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
DB 635 TGCAGATCGTGTGTGACAGAACGAGCCATACACAGCAGCATATTCGGCGAGATTC 694
QY 161 AspSerSerLysAsnIlePheLeuGlyGluLysAlaAlaLysTrpLysAsnProAspGly 180
DB 695 GATTCCTCCAAAATATCTTCTTGAGAGAAAGACAAATGAGAAAACCCGATGGA 754
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProAspGlyGlyPheThr 200
DB 755 CACATGATGACACTACCAACGAGTGTCTGTGTGTGACACCGCAGAGAGGGCTTACCC 814
QY 201 GluGlnSerGlnProGlyValITPARGluIleSerValCysGlyAspValTyrThrLeu 220
DB 815 GAGGAATCCAGGCTGAGTGTGAGAGAGATCTCTGTGTGGAGATGATACACCTTG 874
QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
DB 875 CGAGAGACCAAGTCCGCGCCAGAGAGGAGAAAGTGTGGAATGAGACCAAGTCTCG 934
QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTyrPargThrAlaAspGly 260
DB 935 CAAAGACGGCTCCCTCATTTGACTGTGTGGGCGCATCTCTCTGGAACACCGCAGATGGC 994
QY 261 LeuPheHisThrProThrGlnLysValHisIleGluLysAlaLeuArgGlnIleAsnAla 280
DB 995 CTTTTCAGGCTCTACTCAGAGACATAGAAAGCCCTCGGCGAGAGATCAATGACGCC 1054

```

```

QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgGlyGlu 300
DB 1055 CGACCCAGTGGCCCGGGGCTTAAACCTCGCTTCCCGACATCAACCGGAGAAA 1114
QY 301 ValValGluGluLysGlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
DB 1115 GTGTGGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1174
QY 321 AsnTyrGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
DB 1175 CACTGGGGCCATGAGAGCGACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1234
QY 341 ValGlyProTyrAlaProLeuTyrLeuGlyCysGluAlaGlyPheTyrValAlaPheAla 360
DB 1235 GTGGGCGCTTACCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1294
QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr 380
DB 1295 NNNNNNNCTATCTTTCACCCCTCGGGACGCTCTTTCAGAGAGAGAGAGAGAGAGAGAG 1354
QY 381 TyrSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaLysProPheCys 400
DB 1355 TGGTCCAGATCCCACTGCGCCAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1414
QY 401 AlaThrGlnLeuValGlyGlnGlnAsnGlyIleLysLeuIlePheGlnGlyProIleAsp 420
DB 1415 GCCACGACGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1474

RESULT 10
ABL58447
ID ABL58447 standard; DNA; 1257 BP.
XX
AC ABL58447;
XX
DT 30-JUL-2002 (first entry)
XX
DE Human pellino-1 polypeptide coding sequence.
XX
KW Nuclear factor-kappaB; NF-kappaB; p38; pellino; interleukin-1; virucide;
KW antibacterial; fungicide; protozoicide; antischistosomal; antirheumatic;
KW antirheumatic; antiinflammatory; antithrombotic; antineoplastic; neuroprotective;
KW neurotropic; antitumor; human; pellino-1; gene; ds.
XX
OS Homo sapiens.
XX
FH Key 1.1257 location/Qualifiers
FT CDS /tag= a
FT /product= "pellino-1"
XX
XX WO200183739-A2.
XX
XX 08-NOV-2001.
XX
XX 27-APR-2001; 2001WO-US013676.
XX
XX 28-APR-2000; 2000US-0200198P.
XX
XX (IMMUNEX CORP.
XX
XX Bird TA, Cosman DJ;
XX
XX WPI; 2002-066532/09.
XX
XX P-PDB; ABB07920.
XX
XX New pellino polypeptides for identifying compounds that alter polypeptide
XX activity, treating pathogenic infection or inhibiting apoptosis, are
XX capable of stimulating nuclear factor-kappaB- or p38-dependent
XX transcription.
XX
XX Claim 3; Page 56-57; 70pp; English.
XX

```


XX ABX05095;
 AC 17-JAN-2003 (first entry)
 DT Human novel polynucleotide #110.
 DE
 XX Human; gene; ser; genetic disorder; gene mapping; medical imaging; cancer;
 XX neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
 KM Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
 KM osteoarthritis; periodontal disease; liver fibrosis; viral infection;
 KM fungal infection; bacterial infection; autoimmune disease; diabetes;
 KM atopic dermatitis.
 XX Homo sapiens.
 OS
 XX MO200274961-A1.
 PN
 XX 26-SEP-2002.
 PD
 XX 14-MAR-2002; 2002MO-US005109.
 PF
 XX 15-MAR-2001; 2001US-00810173.
 PR
 XX (HYSE-) HYSEQ INC.
 PA
 XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F;
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M;
 PI Wehrman T, Wang J, Wang D, Drmanac RT;
 XX WPI; 2003-040556/03.
 DR P-PSDB; ABU00017.
 XX
 PT New isolated polypeptides and polynucleotides, useful for preventing,
 PT treating or ameliorating medical conditions, such as cancer,
 PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
 PT disorders, and infections.
 XX
 PS Claim 1; SEQ ID NO 110; 235pp; English.
 XX
 CC The invention relates to human polynucleotides and the polypeptides they
 CC encode. The polynucleotides and polypeptides are useful in diagnostics,
 CC forensic, gene mapping, medical imaging, identification of mutations,
 CC responsible for genetic disorders or other traits, assessing biodiversity
 CC and producing many other types of data and products dependent on DNA and
 CC amino acid sequences. They are also useful for preventing, treating or
 CC ameliorating medical conditions, such as cancer, neurodegenerative
 CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
 CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
 CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
 CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
 CC Sequences ABX04986-ABX05511 represent human polynucleotides of the
 CC invention. Note: The sequence data for this patent is not represented in
 CC the printed specification but is based on sequence information supplied
 CC by the European Patent Office
 XX
 SO Sequence 1304 BP; 385 A; 270 C; 305 G; 344 T; 0 U; 0 Other;
 XX
 Alignment Scores:
 Pred. No.: 9.52e-166 Length: 1304
 Score: 1917.00 Matches: 342
 Percent Similarity: 90.5% Conservative: 38
 Best Local Similarity: 81.4% Mismatches: 38
 Query Match: 83.7% Indels: 2
 DB: Gaps: 1
 US-10-041-030-4 (1-420) x ABX05095 (1-1304)
 QY 1 MetPheSerProGluGlnGluGlnHisCysAlaProAlaGluProValIleGlyTyrGly 20
 DB 1 ATGTTTCTCTCTGATCAAGAAATCAT-----CCATCTAAAGCACAAGAAATATATGGT 54
 QY 21 GluLeuValIleGluGlyTyrAsnGlyAlaLeuProAsnGlyAspArgGlyArgGly 40

DB 55 GAACTCATTTCTTACGGTATATAGGCTCTCCCAATAGCGATAGAGAGAGAGAA 114
 QY 41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValIlePheProSerThrValHis 60
 DB 115 AGTAGGTTGGCTTGTATTAAGACCTTAAGCAAAAGGGGTGAAGCCGACCTGTGCAT 174
 QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLeuGlyGlnHisSerIleSer 80
 DB 175 ATTCCTTGATCTCTCAGGCTGCAAGGCAATAGCAAAAGACCAAGCATATGATATCA 234
 QY 81 TyrThrLeuSerArgAsnGlnThrValValGluTyrThrHisAspLysAspThrAsp 100
 DB 235 TATACCTTGTCTCGGGCCAGACTGTGTGTGTGAATATATCATATACAGCAACACAGAT 294
 QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValIleThrAspThrIle 120
 DB 295 ATGTTTCAGATTGGCCGGTGTGACTGAAGCCCATTTGATTTGTATGACTGACGGGT 354
 QY 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 355 CCGGAAGTCAAGTAAATTTCTGATACAGTCAGTACAAAGCATATATCAAGATTGGC 414
 QY 141 CysArgIleValCysAspArgAsnGluProTyrThrAlaArgIlePheAlaAlaGlyPhe 160
 DB 415 TGCAGATCATATGTGACAGCAATCCCTTTACAGACGGATTTATGCTGCAGGATTT 474
 QY 161 AspSerSerLysAsnIlePheLeuGlyGluValAlaIleTyrTrpLysAsnProAspGly 180
 DB 475 GACTCATCAAAAACATCTTTCTGGGAGAGAGCGTCGCAATGAGAAAGACATCAGATGGA 534
 QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleValMetHisAspArgGlyGlyPheThr 200
 DB 535 CAGATGATGGCTTGAACACTAATGCTTCTGTGATGATGATCAGCAAGTGGTTCACA 594
 QY 201 GluGluSerGlnProGlyValIleTyrArgGluIleSerValCysGlyAspValIleThrLeu 220
 DB 595 GAAGACTCCAGCGCTGGAATATGAGAGAAATATCGGTGTGGAATGATATTACCTTA 654
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyValLeuValGluSerGluThrAsnValIleu 240
 DB 655 CGTGAACCGCATGCGCTCGACGAGAGAAATAGTGAATTTAAACCAATCACTTA 714
 QY 241 GluAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuTyrArgThrAlaAspGly 260
 DB 715 CAAAGTGGCTCGTAAATTAATCTGTGTGCAACATTGTATGGGCTATCGAGAAGGC 774
 QY 261 LeuPheHisThrProThrGlnIleValIleGluAlaLeuArgGlnGluIleAsnAlaIle 280
 DB 775 CTTTCCACACTCTCTACCGCGAAGCATTTTGAAGCTTTAAGACAGAAATCAATGCAGCA 834
 QY 281 ArgProGlnCysArgProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300
 DB 835 CGACCTCAGAGCCCTGTAGGGTTCAACACATGCAATTTCTCTGTATGAAGAGAAAGAAC 894
 QY 301 ValValGluGluGlyGlnProTyrAlaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
 DB 895 GTTGATGATAAACCAACCATGAGTATATCTAACTGCGGCCATGTATCATGCTATCAT 954
 QY 321 AsnTyrGlyHisArgSerAspThrGlnAlaAsnGluArgGlyCysArgProMetCysArgThr 340
 DB 955 AACCTGGGAACAAAGAAAGAAAGCTGATGAAAGATCGTGAATGCTCTATGTAGGCTT 1014
 QY 341 ValGlyProTyrValProLeuTyrPheGlyCysGlnAlaGlyIleThrValAlaPheAlaGly 360
 DB 1015 GTTGTGCTCCATATGCTCTGTGTGTGATGTAAGCTGATTTATGTGACGCGGCG 1074
 QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlyLysSerAlaLysTyr 380
 DB 1075 CTTCCAAACCATCGTTTACCCGCTGTGAGCATGTGTGTTCAAGAAAGACAACTGCTAT 1134
 QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysArgProPheCys 400

DB 1135 TGGTCCAGATCCCACTTCATGTAATCTTTCATGACGCTGTCCTTTGT 1194
QY 401 AlAThGlnleuValGlyGlnIleuAsnCylleuLeuLeuPheGlnGlyProIleAsp 420
DB 1195 GCACATCATGTTGGCTGGTGAACAGGCTACATCAGACTTATTTTTCAGAGACCTCTAGAC 1254
RESULT 13
AAL46320
ID AAL46320 standard; cDNA; 3222 BP.
XX AAL46320;
XX
DT 19-JUL-2002 (first entry)
XX
DE Human M30 variant D coding sequence SEQ ID NO: 9.
XX
XX Neurodegenerative disease; M30; M31; M32; M33; stroke;
KW fragile X syndrome; Huntington's disease; Parkinson's disease;
KW Alzheimer's disease; multiple sclerosis; ovarian cancer;
KW neurodegeneration; immune disorder; autoimmune disease; allergy;
KW infection; leukaemia; inflammation; neuroprotective; cerebroprotective;
KW immunosuppressive; cytostatic; nootropic; antiparkinsonian; antiallergic;
KW vaccine; antiinflammatory; gene; ss.
XX
XX Homo sapiens.
XX
XX MO200221138-A2.
XX
XX 14-MAR-2002.
XX
XX 07-SEP-2001; 2001WO-BP010366.
XX
XX 07-SEP-2000; 2000US-00657479.
XX
XX (AXAR-) AXARON BIOSCIENCE AG.
XX
XX PA
PI Schneider A, Hiemisch H, Rosner M, Klugmann M, Naim J;
PI Eisenhardt G, Kuner R, Lanahan A, Worley P, Spielvogel D, Scheek S;
XX
XX WPI; 2002-292287/33.
DR P-PSDB; AAO17519.
XX
XX
PT Diagnosis of neurodegenerative disease comprises detecting level of M30-
PT family proteins.
XX
XX
XX Claim 3; Page 97-101; 130pp; German.
XX
XX The present invention relates to a method of diagnosing neurodegenerative
XX diseases, comprising determining the concentration of a protein in a body
XX sample, where the protein may be M30 or a variant thereof, M31, M32 or
XX M33. The method is used to diagnose neurodegenerative diseases,
XX particularly stroke but also e.g. fragile X syndrome, Huntington's,
XX Parkinson's and Alzheimer's diseases, multiple sclerosis etc. Also
XX overexpression of M31 can be used for diagnosis of carcinoma and sarcoma.
XX especially ovarian cancer. The proteins can be used to identify specific
XX ligands, potentially useful for treating neurodegeneration, immune-system
XX disorders (e.g. autoimmune diseases, allergy, viral infection, leukaemia,
XX inflammation etc.), carcinoma and sarcoma. Inhibitors of the interaction
XX between the proteins and the protein kinase IRAK-1 can be used to treat
XX neurodegeneration. The present sequence is a coding sequence of a protein
XX used in the method of the invention
XX
SQ Sequence 3222 BP; 980 A; 576 C; 641 G; 1025 T; 0 U; 0 Other;
XX
Alignment Scores:
Pred. No.: 3,27e-165 Length: 3222
Score: 1917.00 Matches: 342
Percent Similarity: 90.5% Conservative: 38
Best Local Similarity: 81.4% Mismatches: 38
Query Match: 83.7% Indels: 2
DB: Gaps: 1
US-10-041-030-4 (1-420) x AAL46320 (1-3222)

QY 1 MetPheSerProGlyGlnGlnIleuHicCyAlaProAsnLysGluProValLysTYRGLY 20
DB 104 ATGTTTCTCTCATCAAGAAATCAT-----CACTTAAACACACAGTAAATATGCT 157
QY 21 GlnuLeuValLleuGlyTYRANGLYALaleuProAsnGlyAspArgGlyArgArgLys 40
DB 158 GAATCATTTCTTAGGGTATATGGGTCTCTCCAAATGGCATATAGAGAAAGAGAAA 217
QY 41 SerArgPheAlaleuTYRlySArgProLYAlaAsnGlyValLysProSerThrValHis 60
DB 218 AGTAGTTTCTGTTTAAAGACCTAAGGCAATGGGGTGAAGCCACAGCAGTGCAT 277
QY 61 ValLleSerThrProGlnLleSerLysAlaLleSerCyLysGlyGlnHisSerLleSer 80
DB 278 ATGCTTGATCTCTCAGGCTGCAAGGCAATAGCAACAAAGACAGCATGATATCA 337
QY 81 TYRThrLleuSerArgAsnGlnThrValValGlyTYRHisAspLysAspThrAsp 100
DB 338 TATACCTTGTCTGGGCCGAGACTGTGTGTGTTGAATATCTCATGACAGCAACAGAT 397
QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValIleAspThrIle 120
DB 398 ATGTTTCAGATGGCCGGTCCAGTGAAGCCCATTTGATTTGATGACACCGCT 457
QY 121 SerGlySerGlnAsnThrAspGlnLleGlnLleThrGlnSerThrLleSerArgPheAla 140
DB 458 CTTGAAGTCAAGATCTGATCTGATACAGTCAATCAAGCACTATATCAAGTTTGGC 517
QY 141 CyArgGlyValCyAspArgAsnGluProTYRThrAlaArgLlePheAlaArgLysPhe 160
DB 518 TCGAATATCATATGTGAACGGAATCTCTTACAGCAGATTTATGTGTCAGATTT 577
QY 161 AspSerSerLysAsnLlePheLeuGlyLysAlaAlaLysTrpLysAsnProAspGly 180
DB 578 GACTCATCAAAAACATCTTTTGGGAGAAAGGTGCCAAATGAAAGATCATGAGATGA 637
QY 181 HisMetAspGlyLeuThrThrAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
DB 638 CAGATGATGGCTTGACCACTAATGTGTCTGTATGATCATCAGCAATGGGTTCACA 697
QY 201 GlnGluSerGlnProGlyValTrpArgGlnLleSerValCyGlyAspValTYRThrLleu 220
DB 698 GAAGACTCCAGGCTGGAATATGAGAAATATCGGTGTGAAATGATTTAGCTTA 757
QY 221 ArgLuthrArgSerAlaGlnArgLysLysLeuValGluSerGlnThrAsnValLleu 240
DB 758 CGTGAACCAAGATCGGCTCAGAGAGAGAAATGTGAAATTCAGATCACTTA 817
QY 241 GlnAspGlySerLleuLleAspLysGlyValThrLeuLeuTrpArgLysAlaAspGly 260
DB 818 CAAGATGGCTCGTAAATGACCTCTGTGGCAACATGTTATGGGTGATCTGCAAGAGC 877
QY 261 LeuPheHisThrProThrGlnLysHisLleGlnAlaLeuArgGlnLleAsnAlaAla 280
DB 878 CTTTCCACACTCTCAACGGTGAACATTTAGAGCTTTAGAGAGAAATCATATGACGA 937
QY 281 ArgProGlnCyAspProValGlyLeuAsnThrLleuAlaPheProSerLleAsnArgLysGlu 300
DB 938 CGACCTCAGGCTCTGTAAGGTTCAACACATGCACTTCTAGTATGAAGAGAAAGAC 997
QY 301 ValValGlnGlnLysGlnProTrpAlaTYRLeuSerCyGlyHisValHisGlyTYRHis 320
DB 998 GTTGTATGATGAAAAACACCATGCGTATATCTAATCGCGCATATGATCAT 1057
QY 321 AsnTrpGlyHisArgSerAspThrGlnLleAsnGluArgGluCyAspProMetCyAspThr 340
DB 1058 AACTGGGAAACAAAGAAAGACCTGATGAAAGATCGAATGTCTATGTATAGTCT 1117
QY 341 ValGlyProTYRValProLeuTrpLeuGlyCyGlnAlaGlyPheTYRValAspAlaGly 360
DB 1118 GTTGTCCCATTTCTCTGTGTGCTTGGATGTGAAGTGAATTTATATGAGACGCGGC 1177

QY 361 ProProThriAlaPheThrProCyseGlyValAlaCyseSerGluYseAlaIleYr 380
Db 1178 CTTCCAAACCATGGTTTACCGGTGGGCAATGTGTGTTCAAGAAAAGACACGCTAT 1237
QY 381 TrpSerGlnIleProLeuProHleGlyThrHisAlaPheHisAlaAlaCyseProPheCyse 400
Db 1238 TGGTCCAGATCCCACTTCTCATGTGATCATATCTTTCATGAGCTCTCCCTTTTGT 1297
QY 401 AlaThrGlnLeuValGlyGluGlnAseCyseAlleYseIlePheGlnGlyProIleAsp 420
Db 1298 GCACATCAGTTGGCTGGTGAACAGGCTCATCATCACTTATTTTCAAGGACCTTAGAC 1357
RESULT 14
ID AAL46319 standard; cDNA; 3526 BP.
AC AAL46319;
XX
XX 19-JUL-2002 (first entry)
DE Human M30 variant C coding sequence SEQ ID NO: 7.
XX
XX Neurodegenerative disease; M30; M31; M32; M33; stroke;
KW Fragile X syndrome; Huntington's disease; Parkinson's disease;
KW Alzheimer's disease; multiple sclerosis; ovarian cancer;
KW neurodegeneration; immune disorder; autoimmune disease; allergy;
KW infection; leukaemia; inflammation; neuroprotective; cerebroprotective;
KW immunosuppressive; cycostatic; mototropic; antiparkinsonian; anti-allergic;
KW vitruide; anti-inflammatory; gene; ss.
XX
OS Homo sapiens.
XX
PN WO200221138-A2.
PD 14-MAR-2002.
XX
XX 07-SEP-2001; 2001WO-EP010366.
XX
XX 07-SEP-2000; 2000US-00657479.
XX
XX (AXAR-) AXARON BIOSCIENCE AG.
XX
XX Schneider A, Hiemisch H, Rosner M, Klugmann M, Naim J;
PI Eisenhardt G, Kuner R, Lanahan A, Worley P, Spielvogel D, Scheek S;
XX WPI; 2002-292287/33.
DR P-PSDB; AAO17518.
XX
XX Diagnosis of neurodegenerative disease comprises detecting level of M30-
PT family proteins.
XX
PS Claim 3, Page 91-95; 130pp; German.
XX
XX The present invention relates to a method of diagnosing neurodegenerative
CC diseases, comprising determining the concentration of a protein in a body
CC sample, where the protein may be M30 or a variant thereof, M31, M32 or
CC M33. The method is used to diagnose neurodegenerative diseases;
CC particularly stroke but also e.g. Fragile X syndrome, Huntington's,
CC Parkinson's and Alzheimer's diseases, multiple sclerosis etc. Also
CC overexpression of M31 can be used for diagnosis of carcinoma and sarcoma,
CC especially ovarian cancer. The proteins can be used to identify specific
CC ligands, potentially useful for treating neurodegeneration, immune-system
CC disorders (e.g. autoimmune diseases, allergy, viral infection, leukaemia,
CC inflammation etc.), carcinoma and sarcoma. Inhibitors of the interaction
CC between the proteins and the protein kinase IRAK-1 can be used to treat
CC neurodegeneration. The present sequence is a coding sequence of a protein
CC used in the method of the invention
XX
SQ Sequence 3526 BP; 1076 A; 618 C; 685 G; 1147 T; 0 U; 0 Other;

Alignment Scores: 3,7e-165 Length: 3526
Pred. No.: 1917.00 Matches: 342

Percent Similarity: 90.5% Conservative: 38
Best Local Similarity: 81.4% Mismatches: 38
Query Match: 83.7% Indels: 2
DB: 6 Gaps: 1
US-10-041-030-4 (1-420) x AAL46319 (1-3526)
QY 1 MetPheSerProGlyGlnGluGlnHisCyseAlaProAsnIleGluProValIleYrGly 20
Db 408 ATGTTTCTTCCTGATCAAGAAATCAT-----CAATCTAAAGCACAGTAAATATGGT 461
QY 21 GluLeuValIleGlyIleYrAengIleAlaLeuProAsnGlyAspArgGlyArgGly 40
Db 462 GAATCATTTGCTTAAAGGATATATAGGTCTCTCCAAATGGCCATAGAGAGAGAGAAA 521
QY 41 SerArgPheAlaLeuYrIleYrAspArgProValAlaAsnGlyValIleProSerThrValHis 60
Db 522 AGTAGGTTGGCTTGTAAAGACCTTAAGGCAAAATGGGTGAAGCCAGACATGTGCAT 581
QY 61 ValIleSerThrProGlnAlaSerIleYrAlaIleSerCyseIleGlyGlnHisSerIleSer 80
Db 582 ATTCCTTGTACTCTCAGGCTGCAAGGCAATAGCAAAAGACATGACATGACATATCA 641
QY 81 TyrThrLeuSerArgAsnGlnThrValIleValIleGlyIleThrHisAspYrAspThrAsp 100
Db 642 TATACCTTGTCTGGGCCAGACTGTGTGTGAATATATCATGACAGACACAGAT 701
QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValIleThrAspThrIle 120
Db 702 ATGTTTCAGATTGGCCGGTGCAGTGAAGCCCATTTGATTTGTATACATGACAGCGTT 761
QY 121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 762 CTTGAAGTCAAGATATTTCTGATACAGCTCATGACAAAGCATATATCAAGATTTGCC 821
QY 141 CysArgIleValCysAspArgAsnGluProYrThrAlaArgIlePheAlaAlaGlyPhe 160
Db 822 TGCAGATCATATGTGAAGCAATCTCCCTTACAGCAGGATTAATGCTGACAGATTT 881
QY 161 AspSerSerIleAsnIlePheLeuGlyGluValAlaAlaIleThrIleYrAspProAspGly 180
Db 882 GACTCATCAAAAACATCTTCTTGGGAGAGAGGCTCCAAATGGAAGACATCAGATGGA 941
QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleValIleMetHisProArgGlyGlyPheThr 200
Db 942 CAGATGATGGCTTGAACCACTAATGCTTCTTGTATGATCATCAGCAATGAGTTCA 1001
QY 201 GluGluSerGlnProGlyValIleTPArgGluIleSerValCyseGlyAspValYrThrLeu 220
Db 1002 GAAGACTCCAGGCTGGAATATGAGAGAAATATCGGTGTGGAATATGATTTAGCTTA 1061
QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyValLeuValGluSerGluThrAsnValLeu 240
Db 1062 CGTGAACCAAGATCGCTCAGCAGAGAGGAAATATGTGGAATTTAAMCCATTCAGTTA 1121
QY 241 GluAspGlySerIleLeuAspLeuCyseGlyAlaThrLeuLeuThrProThrAlaAspGly 260
Db 1122 CAGAGATGCTCGTTAATGACCTCTGTGTGCAACATGTTATAGGCTATGCAAGAGGC 1181
QY 261 LeuPheHisThrProThrGlnIleValHisIleGluAlaLeuArgGlnIleAsnAlaAla 280
Db 1182 CTTTCCACACTCTACCCGTAAGCATTTTGAAGCTTTTGAAGAGAAATCATATGACGCA 1241
QY 281 ArgProGlnCyseProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgGlyGlu 300
Db 1242 CGACCTCAGGCTCTAGAGGTTCAACACACTGCACTTCTCTAGTATGAAGAGAAAGAC 1301
QY 301 ValValGluGluGluGlnProThrAlaIleYrLeuSerCyseGlyIleValHisGlyThrHis 320
Db 1302 GTTGTAGATGAATAAACACATGCTATATCTAACTGCGGCAATGATGACATGCTATCAT 1361
QY 321 AsnThrGlyHisArgSerAspThrGluAlaAsnGluArgGluCyseProMetCyseArgThr 340

Db 1362 AACGGGAAACAAAGAGACGATGAGAAAGATGTAATGCTTATGTAGGTC 1421
 QY 341 VALGlyProTyrValProLeuTTPLeuGlyCysGluAlaGlyPheTyrValAlaPalaGly 360
 Db 1422 GTTGTCCTCATGCTCTGCTGCTGATGAGATGATTTATATGAGACGCGGC 1481
 QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTyr 380
 Db 1482 CCCCACACCATGCTGTTAGCCCTGCTGGCATGTGTTCACAAAAGACACTGCTAT 1541
 QY 381 TRPSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaLysProPheCys 400
 Db 1542 TGTGTCCAGATCCCATCTCTCATGAGTACATATTTCATGACGCTGCTTTGT 1601
 QY 401 AlaThrGlnLeuValGlyGlnHisCysIleLysLeuIlePheGlnGlyProIleAsp 420
 Db 1602 GCACATCATGTGCTGTGTAACAAGCTACATCAGACTATTTTTCAGAGACCTTAGAC 1661
 RESULT 15
 ID ADF83102 standard; cDNA; 7136 BP.
 AC ADF83102;
 XX ADF83102;
 DT 26-FEB-2004 (first entry)
 XX
 DE Human Pellino homologue 1 gene, overexpressed in cancer.
 KM Human; Pellino homologue 1; Pellinol; cancer; cytostratic; vaccine;
 XX gene therapy; gene; oncogene; ss.
 OS Homo sapiens.
 XX
 FH Key Location/Qualifiers
 FT CDS 4039..5295
 FT /*tag= a
 XX
 PN MO2003100000-A2.
 XX
 PD 04-DEC-2003.
 XX
 PF 22-MAY-2003; 2003MO-US016049.
 XX
 PR 24-MAY-2002; 2002US-0382606P.
 XX
 PR 25-JUL-2002; 2002US-0398099P.
 XX
 PA (TULA-) TULARIK INC.
 XX
 PI Li J, Ma D, Yang J;
 XX
 XX WPI; 2004-035118/03.
 DR P-PSDB; ADF83103.
 DR GENBANK; NM_020651.
 XX
 PT Diagnosing a cancer in a mammal comprises determining RecQL5, CTKL,
 PT USP1, MCL1, or Pellino 1 gene copy number in a biological sample from a
 PT region of the mammal that is suspected to be precancerous or cancerous.
 XX
 PS Claim 179; SEQ ID NO 9; 174bp; English.
 XX
 CC The present sequence is that of the gene encoding human Pellino homologue
 CC 1 (Pellinol), a protein associated with the kinase domain of activated
 CC Pelle. The invention is based on the finding of the overexpression of
 CC Pellinol and other genes (RecQL5, CTKL, USP13 and MCL1) in certain
 CC cancers, including breast cancer, colon cancer, lung cancer and ovarian
 CC cancer, and the frequent amplification of these genes in cancer cells.
 CC The genes, and their expression products, can be used diagnostically or
 CC as targets for cancer therapy. They can also be used to identify and
 CC design compounds useful in the diagnosis, prevention and therapy of
 CC tumours and cancers, in vaccine development, and in methods for
 CC determining the efficacy of a treatment regime. A claimed method for
 CC inhibiting cancer or precancerous growth, especially in ovarian or lung
 CC tissue, uses an inhibitor that interacts with Pellinol DNA or RNA. The

CC inhibitor is a small interfering RNA (siRNA), microRNA (miRNA), an
 CC antisense RNA, and antisense DNA, a decoy molecule, a decoy DNA, a
 CC ribozyme or small molecule.
 XX
 SQ Sequence 7136 BP; 1945 A; 1333 C; 1399 G; 2459 T; 0 U; 0 Other;
 Alignment Scores:
 Pred. No.: 9 67e-165 Length: 7136
 Score: 1917.00 Matches: 342
 Percent Similarity: 90.5% Conservative: 38
 Best Local Similarity: 81.4% Mismatches: 38
 Query Match: 83.7% Indels: 2
 DB: 12 Gaps: 1
 US-10-041-030-4 (1-420) x ADF83102 (1-7136)
 QY 1 MetPheSerProGlyGlnGluHisCysAlaProAsnLysGluProValLysTyrGly 20
 Db 4039 ATGTTTCTCTGATCAAGAAATCAT-----CATCTAAAGCACCACTAAATATGCT 4092
 QY 21 GluLeuValValLeuGlyTyrAngGlyAlaLeuProAngLysAspArgGlyArgGly 40
 Db 4093 GAACCTATGCTTGAAGCTATATAGGCTCTCCCAATGCGATGAGAGAGAGAA 4152
 QY 41 SerArgPheAlaLeuTyrTyrAspProLysAlaAngLysValLysProSerThrValHis 60
 Db 4153 AGTAGGTTGCTTGTAAAGACTAGGCAAAAGGGGTGAAGCCAGACTGTGCAT 4212
 QY 61 ValIleSerThrProGlnLaseLysLysAlaIleSerCysLysGlyGlnHisSerIleSer 80
 Db 4213 ATTGCTGTACTCTCTAGGCTGCAAGGCAATAGACAAAGACACAGCATGACATCA 4272
 QY 81 TyrThrLeuSerArgAngGlnThrValValGluTyrThrHisAspLysAspThrAsp 100
 Db 4273 TAACTTTGCTCGGGCCAGACTGTGGTGAATATACATACAGACAGACAGAT 4332
 QY 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
 Db 4333 ATGTTTCAGATTGGCCGGTGAAGCTGAAAGCCCATTTGATGTAACAGCGGT 4392
 QY 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 Db 4393 CCTGAAGTCAAGATTAATCTGATACAGTCAAGTCAAGACATATACAGATTGGC 4452
 QY 141 CysArgIleValCysAspArgAngLysProTyrThrAlaArgIlePheAlaGlyPhe 160
 Db 4453 TGCAGATCATATGTAAGCAGATCTCTTACAGCAGGATTTACTGCGAGATT 4512
 QY 161 AspSerSerLysAsnIlePheLeuGlyGlyValAlaLysTyrLysAsnProAspGly 180
 Db 4513 GACTCATCAAAAACATCTTTCTGGGAGAAAGGCTGCCAAATGGAAGCATCAGATGGA 4572
 QY 181 HisMetAspGlyLeuThrThrAngGlyValLeuValMetHisProArgLysGlyPheThr 200
 Db 4573 CAGATGATGGCTTGAACCATATGCTGTCTTGATCATACACGCAATGAGTTCAAC 4632
 QY 201 GluGluSerGlnProGlyValTyrPargGluIleSerValCysGlyAspValTyrThrLeu 220
 Db 4633 GAAAGATCCAGCTGAGATATGAGAGAAATATGCTGTGGAATATATTAGCCTA 4692
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
 Db 4693 CGGAAACACAGATCGGCTCAGACAGAGCAAAATGCTGGAATGGAACCATCATGTTA 4752
 QY 241 GlnAspGlySerLeuIleAspLeuCysGlyAlaThrLeuLeuThrArgThrAlaAspGly 260
 Db 4753 CAAGATGCTCTGTAATTATACCTGTGTGGAACATTTATGGGTACTGCAAGAGC 4812
 QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAla 280
 Db 4813 CTTTCCACACCTCTTACCGTGAAGCATTTTGAAGAGCTTTTGAAGCAAGAAATCATGAGCA 4872
 QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGlu 300

```
Db      ||||| 4873 CGACCTCAGGCGCTGAGGGTTCAACACACTAGCATTTCTAGTATGAAGAGAAAGAC 4932
Qy      ||||| 301 ValValGluGluIuyluysGlnProTPrAlaTyrIeuSerCysGlyHisValHisGlyTyrHis 320
Db      ||||| 4933 GTGTGATGATAAACAACACATGGGTATATCTAACTGCGGCCCATGTACATGGCTATCAT 4992
Qy      ||||| 321 AsnTyrGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
Db      ||||| 4993 AACTGGGGAACAACAAGAAACGATGAGAAAAGATGTGAATGTCTATGTAGTCT 5052
Qy      ||||| 341 ValGlyProTyrValProIeuTyrPleuGlyCysGluAlaGlyPheTyrValAspAlaGly 360
Db      ||||| 5053 GTTGGTCCCTATGTCTCTGTGGCTTGATGATGAAGCTGGATTGTATGTGACGCGGC 5112
Qy      ||||| 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluIuysSerAlaLysTyr 380
Db      ||||| 5113 CTTCAACCCATGCGTTAGCCCGTGGGCATGTGTTCAGAAAAGACCAACTGCCTAT 5172
Qy      ||||| 381 TrpSerGlnIlePProIeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
Db      ||||| 5173 TGGTCCAGATCCCACTTCTCATGTACTCATCTTTTCATGAGGCTGTCCCTTTGT 5232
Qy      ||||| 401 AlaThrGlnIeuValGlyGluGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
Db      ||||| 5233 GCACATCAGTTGGCTGGTGAACAAGGCTACATCAGACTTATTTTCAAGGACCTTAGAC 5292
```

Search completed: March 2, 2006, 05:14:48
Job time : 823 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Biocollaboration Ltd.

OM protein - nucleic search, **us-10-041-030-4**

Run on: March 2, 2006, 05:01:06 ; Search time 6947 Seconds
(without alignments) 3436.627 Million cell updates/sec

3436.627 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MESPGEHRCAPKPKVYKX.....ATQLVGEQNCIKIFQSPID 420

Scoring table:
BLOSUM62
Xgapop 10.0, Xgapext 0.5
Ygapop 10.0, Ygapext 0.5
Fgapop 6.0, Fgapext 7.0
Delop 6.0, Delcxt 7.0

Searched: 5883141 seqs, 28421725653 residues

Total number of hits satisfying chosen parameters: 11766282

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-Processing: Minimum Match 0%

Maximum Match 100%
Listing first 45 summaries

Command line parameters:
-MODE=frame+ p2n.model -DEV=x1h
-Q=/abs/ABSSMBR.epool/US10041030/runat_01032006_134404_21821/app_query.fasta_1
-DB=GenEmbl -QFMT=fasta -SUFFIX=p2n.rge -MINMATCH=0.1 -LOOPEXT=0 -LOOPEXT=0
-UNITS=bits -START=1 -END=1 -MATRIX=blomsum62 -TRANS=human40.cdd -LIST=45
-DOCALIGN=200 -THR SCORE=pct -THR MAX=100 -THR MIN=0 -ALIGN=15 -MODE=LOCAL
-OUTFMT=pro -NORM=ext -HEAPSIZE=500 -MINLEN=0 -MAXLEN=2000000000 -HOST=abss808
-USER=US10041030 @CN 1_1 4939 @runat_01032006_134404_21821 -NCPU=6 -ICPU=3
-NO MMAP -NEG SCORES=0 -WAIT -DSELOCK=100 -LONGLOG -DEV TIMEOUT=120
-WARN TIMEOUT=30 -THREADS=1 -XGAPOP=10 -XGAPEXT=0.5 -FGAPOP=6 -FGAPEXT=7
-YGAPOP=10 -YGAPEXT=0.5 -DELOP=6 -DELEXT=7

Database :

GenEmbl: *
1: gb ba: *
2: gb in: *
3: gb env: *
4: gb om: *
5: gb ov: *
6: gb pac: *
7: gb ph: *
8: gb pr: *
9: gb ro: *
10: gb sts: *
11: gb sy: *
12: gb un: *
13: gb vi: *
14: gb hcg: *
15: gb pl: *

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match Length | DB ID | Description |
|------------|-------|--------------------|--------|--------------------|
| 1 | 2290 | 100.0 | 1565 6 | AX817203 Sequence |
| 2 | 2290 | 100.0 | 1811 8 | BC009476 Homo sapi |
| 3 | 2290 | 100.0 | 5579 6 | BD190328 Elkt phos |

| RESULT 1 | LOCUS | DEFINITION | ACCESSION | VERSION | KEYWORDS | SOURCE | ORGANISM | REFERENCE | AUTHORS | TITLE | JOURNAL | FEATURES | CDS | ALIGNMENTS |
|----------|------------------------------------|------------------------------------|-----------|------------|-------------|--------|----------------------|--|---|--------------------------------------|---------------------------|---|----------------------|--|
| AX817203 | Sequence 15 from Patent WO0221138. | Sequence 15 from Patent WO0221138. | AX817203 | AX817203.1 | GI:39722594 | | Homo sapiens (human) | Schneider, A., Hiemisch, H., Rosener, M., Klugmann, M., Naim, J., Eisenhardt, G., Kuner, R., Lannahan, A., Worley, P., Spielvogel, D. and Scheek, S. | The m30 gene family and the utilization thereof | Patent: WO 0221138-A 15 14-MAR-2002; | Axaron Bioscience AG (DE) | Location/Qualifiers 1. 1565 /organism="Homo sapiens" /mol_type="unassigned DNA" /db_xref="taxon:9606" 51. 1313 /note="unnamed protein product; codierende Bereich der M3 CDNA des Menschen" | M3 CDNA des Menschen | AF302502 Homo sapi AX668405 Sequence AR483568 Sequence AX298079 Sequence AR483567 Sequence AX298077 Sequence BC072891 Xenopus t BC063200 Xenopus t BD190327 Elkt phos AF302504 Mus muscu BC075973 Dario rer AR483566 Sequence AX298075 Sequence AF300987 Homo sapi AX817197 Sequence BC063611 Homo sapi BC050533 Homo sapi AX817195 Sequence BC050479 Homo sapi CS036913 Sequence CS042641 Sequence CS045865 Sequence AX668403 Sequence AF302505 Homo sapi AX817189 Sequence AJ278859 Homo sapi AR483565 Sequence AX298073 Sequence BC016515 Mus muscu AF302503 Mus muscu AJ719997 Gallus ga BC044117 Xenopus t AC128556 Rattus no BC027062 Mus muscu CO178825 Sequence AC109960 Rattus no CQ720116 Sequence AR483571 Sequence AX298083 Sequence |


```

/codon_start=1
/protein_id="CAE84330.1"
/db_xref="GI:39722595"
/translation="MSPQGEHCACPKYKXGELVYLGNALPNDGRKRPAL
LYRPRANGKRPSTVHISTPQASKAISCGQHSISTLSRNTQVVEVTHDQDMF
QVRSSTESPIDVPTDTISGQNTDEAQITQSTISRPACKIVCDRNPYTAI.PAAGF
DSSKNT.FLGEKRAKWKPNRGMDBGLTNNGLVMMPRGSPFESQPGWRI.SVCGDY
TLHETRSACQGRGLVSESTNVLDQSLIDLCATLLWRADGLFHTPTQRI.BALQD
INARPCQPVGLNTLAFPSINRKEVVEKQPMAYLSCGHYGVNMGHSDTANERE
CPMRVTGPPVPLMIGCEAGFYDAGPPTIAFPICGHVSEKSAKXWSQI.PLHGTHA
FHAACPACATQLVGEONCTILITQGPID"

```

ORIGIN

Alignment Scores:

| Pred. No.: | 1,46e-166 | Length: | 1565 |
|------------------------|-----------|---------------|------|
| Score: | 2290.00 | Matches: | 420 |
| Percent Similarity: | 100.0% | Conservative: | 0 |
| Beet Local Similarity: | 100.0% | Mismatches: | 0 |
| Query Match: | 100.0% | Indels: | 0 |
| DB: | 6 | Gaps: | 0 |

US-10-041-030-4 (1-420) x AX817203 (1-1565)

```

QY 1 MecPheSerProGlyngingluhneCyalaProaenlysgIuproValyTyrgly 20
Db 51 ATGTTTCCCTGGCCAGAGAACTGGGCCCCCAATAAGAGCCAGTGAATACGGG 110
QY 21 GluleuValIleuGlyTyraenGlyAlaIeuProaenGlyAepaGlyAaGly 40
Db 111 GACCTGTGTGTCTGGGTCAATAGTGTCTTCAATGAGATGAGAGCGAGAAA 170
QY 41 SerArgPheAlaIeuTyryArgProLyseAlaenGlyValLyseProserThValHis 60
Db 171 AGTGAATTTGCCCTTCAAGCGGCCCAAGCAATGCTGTCAACCAAGCACCTGCAT 230
QY 61 ValIleSerThProGlnAlaSerLyseAlaIleSerCyalyGlyGlnHisSerIleSer 80
Db 231 GTGATATCCAGCCCAAGGATCCAGGCTATGCTGCAAAAGGTCAACAGATATACC 290
QY 81 TyrThIeuSerArgaenGlnThValValIeuTyrrhIaApLySapLySapThAsp 100
Db 291 TACACTTGTTCAGAAATCAGACTGTGTGTGAGTACACATGATAGATACCGAT 350
QY 101 MetPheGlnValGlyArgSerThGlySerProIleAspPheValIleThAspThIle 120
Db 351 ATGTTTAGGTGGCAATCAACGAAAGCCCTTACACTTCTGTTCACAGACACGATT 410
QY 121 SerGlySerGlnaenThraSpGluAlaGlnIleThrGlnSerThIleSerArgPheAla 140
Db 411 TCTGGCAGCCAGAACACGAGACGAAAGCCCAAGTCAACAGAGCACATATCCAGGTTCC 470
QY 141 CyahArgIleValCyabAPaGlaenGluProTyrrhIaIaArgIlePheAlaIleGlyPhe 160
Db 471 TGCAGGTCTGTGCGACAGAAAGAACTTACACAGACAGATATTTGCCGCCGGAATTT 530
QY 161 AspSerSerLyseAnIlePheLeuGlyGluValAlaIalValSTrPySaenProaGly 180
Db 531 GACTCTTCAGAAACATATTTCTTGGAGAAAGCAGCAAGTGAAGAACCCCGAGCGGC 590
QY 181 HisMetAspGlyLeuThrrhAsnGlyValIleValMetHisProArgGlyGlyPheThr 200
Db 591 CACATGATGGGCTCACTAATAGGGCTCTGTGATGATCATCAAGAGGGGGCTTACCC 650
QY 201 GluGluSerGlnProGlyValTyraGluIleSerValCyaglyAapValTyrrhIeu 220
Db 651 GAGGAGTCCCGAGCCCGGGGTCTGGCGGAGATCTCTGTGTGAGATGTATACACCTTG 710
QY 221 ArgGluThraArgSerAlaGlnGlnArgGlyValIleValGlySerGluThraenValIeu 240
Db 711 CGAGAAACCGAGTGGCCAGCAAGAGAAAGCTGGGGAAGTGAAGACCAACGCTCTG 770
QY 241 GlnAspGlySerLeuIleAspLeuCyaglyValaThrIeuLeuTyrrhIaAspGly 260

```

```

Db 771 CAGAGCGGCTCCCTCATGACTGTGTGGGCACTCTCTCTGAGAAACAGATGGG 830
QY 261 LeuPheIsthrProThrgInlyshIleGluAlaIeuArgGlnIleAnalaA 280
Db 831 CTTTTCATATCTCCACTCGAAGACATTAAGACCTCCGGACGAGATTAAGCCGCC 890
QY 281 ArgProGlnCyabProValGlyIleuAnThrIleuAlaPheProSerIleAnArglyGlu 300
Db 891 CGGCTCAGTGTCTGTGGGGCTCAACACCTGGCTTCCCGACATCAACAGAAAGAG 950
QY 301 ValValGluGluIlySerGlnProTyrrhIaIyrrhIeuSerCyaglyValIleGlyTyrrhIs 320
Db 951 GTGGTGAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1010
QY 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGlyAlaGlyCyseProMetCyArgThr 340
Db 1011 AACTGGGGCCATGGAGTACACGAGGCCAAGAGAGAGAGAGAGAGAGAGAGAG 1070
QY 341 ValGlyProTyrrhIaProLeuTyrrhIeuGlyCyaglyAlaGlyPheTyrrhIaAspAlaGly 360
Db 1071 GTGGGGCCCATATGAGCTCTCTGGCTGTGGAGAGAGAGAGAGAGAGAGAGAG 1130
QY 361 ProProThrHisAlaPheThrProCyaglyHisValCyseSerGlySerAlaIyTyrr 380
Db 1131 CCGCCAACTCATCTTCACTCCCTGTGACAGGTGTCTGGAGAGAGAGAGAGAGAG 1190
QY 381 TrpSerGlnIleProIeuProHisGlyTyrrhIaAlaPheHisAlaIaCyseProPheCy 400
Db 1191 TGGTCTCAGATCCCGTGTCTCAATGAACTCAATGATTAAGCTGTGCTTCTGT 1250
QY 401 AlaThrGlnIleuValGlyGluGlnaenCyalyseLeuIlePheGlnGlyProIleAsp 420
Db 1251 GCTACACAGCTGGTGGGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1310

RESULT 2
BC009476 1811 bp mRNA linear PRI 08-MAR-2005
LOCUS Homo sapiens pel1no homolog 2 (Drosophila), mRNA (cDNA clone
DEFINITION MGC:15066 IMAGE:3942712), complete cds.
ACCESSION BC009476
VERSION BC009476.2 GI:33872093
KEYWORDS MGC.
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
Homnidae; Homo.
1 (bases 1 to 1811)
Strausberg RL, Feingold EA, Grouse LH, Derge JG, Klausner RD,
Collins FS, Wagner L, Shenmen CM, Schuler GD, Altschul SF, Zeeberg
B, Buettow KH, Schaefer CF, Bhat NK, Hopkins RF, Jordan H, Moore T,
Max SI, Wang J, Hsieh P, Diatchenko L, Marusina K, Farmer AA, Rubin
GM, Hong L, Stapleton M, Soares MB, Bonaldo MP, Casavant T,
Scheetz TE, Brownstein WJ, Ustin TB, Toshiyuki S, Carninci P,
Prange C, Raha SS, Loggellano NA, Peters GJ, Abramson RD, Mullaby
SJ, Bosak SA, McKern PJ, McKernan PJ, Malek JA, Gunaratne PH,
Richards S, Worley KC, Hale S, Garcia AM, Gay LJ, Hulyk SM,
Villalon DK, Wuzny DM, Sodergren EJ, Lu X, Gibbs RA, Faney J,
Heiton E, Kettelman M, Madan A, Rodrigues S, Sanchez A, Whiting M,
Madan A, Young AC, Shevchenko Y, Bouffard GG, Blakeley RW,
Touchman JW, Green ED, Dickson MC, Rodriguez AC, Grimwood J,
Schmutz J, Myers RM, Butlerfield YS, Krzywicki MI, Skalska U,
Smalins DE, Schnerch A, Schein JE, Jones SJ and Maitra MA.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
JOURNML
PUBMED
2 (bases 1 to 1811)
REFERENCE
Director MGC Project.
Direct Submission
Submitted (25-JUN-2001) National Institutes of Health, Mammalian
Gene Collection (MGC), Cancer Genomics Office, National Cancer

```


LOCUS BD190328 5579 bp DNA linear PAT 17-JUL-2003
 DEFINITION Ekl1 phosphorylation related gene.
 ACCESSION BD190328
 VERSION BD190328.1 GI:3300067
 KEYWORDS WO 03008589-A/68.
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 5579)
 AUTHORS Matsuzaki, O., Matenda, A., Nagano, Y. and Suzuki, N.
 TITLE Ekl1 phosphorylation related gene
 JOURNAL Patent: WO 03008589-A 68 30-JAN-2003;
 ASahi Kasei Corp, OSAW MATSUZAKI, AKIO MATSUDA, YUKIKO NAGANO, NAOMI SUZUKI
 COMMENT OS Homo sapiens (human)
 PN WO 03008589-A/68
 PD 30-JAN-2003
 PR 15-JUL-2002 WO 2002JP007174
 PR 18-JUL-2001 JP 01P 218204,31-AUG-2001 JP 01P 263450 PR
 21-JAN-2002 JP 02P 012176
 PI OSAW MATSUZAKI, AKIO MATSUDA, YUKIKO NAGANO, NAOMI SUZUKI PC
 C12N15/54, C12N9/12, G01N33/50, G01N33/15, C07K16/40, A61K39/385, PC
 A61K31/711
 CC Ekl1 phosphorylation related gene
 FH Key Location/Qualifiers
 FT CDS Location/Qualifiers
 1..5579
 /organism="Homo sapiens"
 /mol_type="genomic DNA"
 /db_xref="taxon:9606"
 ORIGIN
 Alignment Scores:
 Pred. No.: 6,28e-16 Length: 5579
 Score: 2230.00 Matches: 420
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 6 Gaps: 0
 US-10-041-030-4 (1-420) x BD190328 (1-5579)
 QY 1 MetPheSerProGluGlnGluGlnHicGValAProAnlysgLUProVallyrTYrGly 20
 DB 177 ATGTTTCCCTGGCCGAGGAACTGCGCCCAATAGAGGCAAGTAATACGGG 236
 QY 21 GluLeuValAlaLeuGlyTYrAengLYAlaLeuProAnlysgLAPATGGLATGATG 40
 DB 237 GAGCTGGTGGTCTCGGGGTAATGATGCTTTACCAATGAGATAGAGACGAGAA 296
 QY 41 SerArgPheAlaLeuTYrLYsArgProLYsAlaAnGLyVallyrProSerThrValHis 60
 DB 297 AGTAGATTTCCCTCTCAAGCGGCGCAAGGCAATGATGTTCAAACCCAGACGCTCAT 356
 QY 61 ValIleSerThrProGlnAlaSerLYsAlaIleSerCylysgLYGlnHisSerIleSer 80
 DB 357 GTGATATCCACGCCGCCAGGATCCAGGCTATGCTGCACAAAGGTCAACACATATATCC 416
 QY 81 TYrThrLeuSerArgAengLInThrValAlaValGluTYrThrHisAaPlyAaPlyThrAsp 100
 DB 417 TACACTTTGTCAGAGAAATCAAGCTGTGTGTGGATGACACACATGATTAAGATACGAT 476
 QY 101 MetPheGlnValGlyATSerThrGluSerProIleAspPheValValThrAspThrIle 120
 DB 477 ATGTTTCAAGGCGGCAATCAACAGAAAGCCCTATGCACTTCGTGTCAAGACAGAT 536
 QY 121 SerGlySerGlnAnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 537 TCTGGCAGCGAAGACAGGAGCCAGACCATGATCAACAGAGACCATATATCCAGGTGCC 596

QY 141 CyAargIleValCyAaPargaengLUProTYrThrAlaArgIlePheAlaGlyPhe 160
 DB 597 TGCAGATGCTGTGGCAGAGAAATGAACCTTACACAGCAGATATTCGCCGCCGATTT 656
 QY 161 AspSerSerLYsAnIlePheLeuGlyGluLYsAlaAlaLYsTrpLYsAnPProAaPly 180
 DB 657 GACTCTTCCAAAACATATTTCTTGGAGAAAGGCAAGGAGGAAAAACCCGACGGC 716
 QY 181 HiMetAaPlyLeuThrThrAengLYValLeuValMeChIProArgLYGlyPheThr 200
 DB 717 CACATGATGGGCTCACTACTAATGGCGTCTGTGTATGATTCACAGGGGGCTTCACC 776
 QY 201 GluGluSerGlnProGlyValITPATGGLIleSerValCyAglYAspValTYrThrLeu 220
 DB 777 GAGAGATCCAGCCGGGGGTCTGGCGGAGATCTGTGTGTGAGATGTGTACCTTG 836
 QY 221 ArgGluThrArgSerAlaGlnGlnAlaGlyLYsLeuValGluSerGlyThrAnValLeu 240
 DB 837 CGAGAAACCAAGTCGGCCCGCAGCAACAGAAAGCTGTGTGAAGTGAACCAAGTCCTG 896
 QY 241 GluAaPlySerLeuIleAaPlyLeuCyAglYAlaThrLeuLeuTrpArgThrAlaAaPly 260
 DB 897 CAGGACGGCTCCCTCATTAACCTGTGTGGGCCACTCTCTCTGGAGAACAGCATGGG 956
 QY 261 LeuPheHisThrProThrGlnLYsHisIleGluAlaLeuArgGlnGluIleAnAla 280
 DB 957 CTTTTCATATCTCAACTCAGAAACATATGAGACCTCCGGCAGGATTAAGCCGCC 1016
 QY 281 ArgProGlnCyAProValGlyLeuAnThrLeuAlaPheProSerIleAnAlyLYsGlu 300
 DB 1017 CGGCTCAGATGCTCTGTGGGGCTCAACACCTGCTCCCGCATCAACAGAAAGAG 1076
 QY 301 ValValGluGluLYsGlnProTrpAlaTYrLeuSerCyAglYHisValHisGlyTYrHis 320
 DB 1077 GTGTGGAGAGAGAGAGCGCTGGGCTATCTAGTTGTGGCCAGTGACGGGATTCAC 1136
 QY 321 AsnTrpGlyHisArgSerAaPThrGluAlaAengLYaArgLYsCyAProMetCyAArgThr 340
 DB 1137 AACTGGGGCATGTGAGATGACAGGAGCCAAAGAGAGAGTCCCATGTTCAGAGACT 1196
 QY 341 ValGlyProTYrValProLeuTrpLeuGlyCyAglYAlaGlyPheTYrValAaPly 360
 DB 1197 GTGGGCCCTATGCTCTCTGCTGGCTGTGGAGGACGATTTATATGAGACGCGAGA 1256
 QY 361 ProProThrHisAlaPheThrProCyAglYHisValCySerGlyLYsSerAlaLYsTYr 380
 DB 1257 CCGCACTATCTTCACTCCCTGTGACAGCTGTGGAGAGATGTGCAAAATAC 1316
 QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaLYsCyAProPheCy 400
 DB 1317 TGGTTCAGATCCCGTTGGCTCATGGAACTCAAGCATTTCAAGCTGTGCCCTTCTGT 1376
 QY 401 AlaThrGlnLeuValGlyGluAnCyAlaLYsLeuIlePheGlnGlyProIleAaP 420
 DB 1377 GCTACACAGCTGATGGGAGCAAACTGCATCAAAATTAATTTCCAGAGTCAATTGAC 1436
 RESULT 4
 AF302502 5597 bp mRNA linear PRI 05-JUL-2001
 LOCUS AF302502
 DEFINITION Homo sapiens pellino 2 (PELL2) mRNA, complete cds.
 ACCESSION AF302502
 VERSION AF302502.1 GI:10242352
 KEYWORDS
 SOURCE Homo sapiens (human)
 ORGANISM Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae; Homo.
 REFERENCE 1 (bases 1 to 5597)
 AUTHORS Resch, K., Jockusch, H. and Schmitz-John, T.
 TITLE Assignment of homologous genes, PELL1/PELL1 and PELL2/PELL2, for the pelle adaptor protein Pellino to mouse chromosomes 11 and 14 and human chromosomes 2p13.3 and 14q21, respectively, by physical

JOURNAL
 PUBMED
 REFERENCE
 AUTHORS
 JOURNAL
 TITLE

and radiation hybrid mapping
 Cytogenet. Cell Genet. 92 (1-2): 172-174 (2001)
 11306823
 2 (bases 1 to 5597)
 Resch, K., Jockusch, H. and Schmitt-John, T.
 Direct Submision
 Submitted (05-SEP-2000) Developmental Biology and Molecular
 Pathology, University of Bielefeld, Universitaetsstrasse 25,
 Bielefeld, NRW 33615, Germany

FEATURES

Source

Location/Qualifiers
 1..5597

/organism="Homo sapiens"
 /mol_type="mRNA"

/db_xref="taxon:9606"

1..5597

/gene="PEL12"

177..1439

/note="similar to Drosophila melanogaster Pellino"

/codon_start=1

/product="pellino 2"

/protein_id="PAGI5390.1"

/db_xref="GI:10242353"

/translation="MESPQGEHCAPNKEPVKGYELVVLGNGALPNDGRKRKRPFA
 LYRPKANGVKPSTVHYITSPQASKATSKGSHISYLSRNQVVEVTHMDQDMF
 QNGRSTSPIDFVVTDTISSQNTDEAQTOSTISRACTIVCDNRPRTARIAPAF
 DSSKNTFLGRKAKAKNPBGDMDELITNGVLMHPRGFTESQGVWREISVCGDY
 TLRTSRSAQQRKLVESRTVLDGSLIDCGATLWRTADGLFHTPTQRIEALRQ
 INARPCQPVGLMTLGPSSINRKEVVEKQPMWALSCGHVGHYHNMHRDTEANERE
 CPMCRIVGPVPLMLCEAFPGFYDAGPPTAFPTFCGHVCSKSKAKYWSQIPLEGTHA
 FHACPCPAQVLQVGEQNCIKLIPGSPID"

ORIGIN

Alignment Scores:

Pred. No.: 6,3e-166 Length: 5597
 Score: 2290.00 Matches: 420
 Percent Similarity: 100.0% Conservative: 0
 Best Local Similarity: 100.0% Mismatches: 0
 Query Match: 100.0% Indels: 0
 DB: 8 Gaps: 0

US-10-041-030-4 (1-420) x AF302502 (1-5597)

QY 1 MetPheserProGIynglunghicCyAlaProAbnlysgluproVallyrtyrGly 20
 DB 177 ATGTTTCCCTGGCCGAGGAACTGGCCGCCCAATAGAGGCGAATATACGGG 236
 QY 21 GluLeuValValleuGlyTyraNGlyAlaLeuProAbnlysaPaArgIyATGArgLy 40
 DB 237 GAGCTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGTGGT 296
 QY 41 SerArgPheAlaLeuTyraNGlyTyraNGlyAlaLeuProAbnlysaPaArgIyATGArgLy 60
 DB 297 AGTAGATTTCCCTCTCAAGCGCGCCCAAGGCAAAATGTCGCAACCAAGCAGCCGTCAT 356
 QY 61 ValIleSerThrProGlnAlaSerTyraAlaIleSerCyAllysglyGlnIleSerIleSer 80
 DB 357 GTGATATCCACGCCCGAGGATCCAGAGCTATCAGTCGCAAAAGGTCAACACAGATATCC 416
 QY 81 TyThrIleuSerArgaNGlynthValValValGlyTyThrIleSaPlysaPlysaPlysaPly 100
 DB 417 TACCTTTGTCAGAGATCAAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 476
 QY 101 MetPheGlnValGlyArgSerThrGlySerProIleAspPheValValThrAspThrIle 120
 DB 477 ATGTTTCAAGGTGGCAATCAAGAAAGCCCTTATCAGCTTCGTTGTCACAGACAGAT 536
 QY 121 SerGlySerGlnIleThrAspGlyAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
 DB 537 TCTGGCCGACGAGACAGAGAGCCAGATTCACACAGAGACACATATCCAGGTGGCC 596
 QY 141 CyAArgIleValCyAAspArgaNGlyProTyThrIleArgIlePheAlaIleGlyPhe 160

DB 597 TGACGATCGTGTGGCAGAGAAATGAACCTTACACAGCAGGATATTCGCCGCCGATTT 656
 QY 161 AspSerSerIyabnIlePheLeuGlyGlnIyAlaIleTyrrPlysaPProAbnly 180
 DB 657 GACTCTTCCAAAACATATTTCTTGGAGAAAGCGCAAGAGGAGAAACCCCGACGGC 716
 QY 181 HiMeAspGlyLeuThrThrAnGlyValLeuValMeHiSPProArgIyGlyPheThr 200
 DB 717 CACTGATGGGCTCACTACTAATGCGCTCTGTGTATGATCATCCAGGGGCTTAC 776
 QY 201 GluIleuSerGlnProGlyValITPArgIuIleSerValCySglYaaPValTyThrLeu 220
 DB 777 GAGAGTCCAGCCCGGGGTCTGGCCGAGATCTGTGTGTGTGTGTGTGTGTGTGTGTGT 836
 QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLyLeuValIleuValIleuValIleuVal 240
 DB 837 CGAGAAACAGGTCTGGCCCGACAGACAGAGAACTGTGTGAAAGTGAGACCAACGTCCTG 896
 QY 241 GluAspGlySerIleuIleAspLeuCyAGlyAlaThrIleuThrTPArgThrIleAspGly 260
 DB 897 CAGAGCGCTCCCTCATTAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 956
 QY 261 LeuPheHisThrProThrGlnIySHIleGlyAlaLeuArgIuIleAsnAlaIle 280
 DB 957 CTTTTCATATCTCCACTCAAGACATAGAGCCCTCGGACAGAGATTAACGCCGCC 1016
 QY 281 ArgProGlnCySProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLyGly 300
 DB 1017 CGGCTCAGTGTCTGTGGGGCTCAACACCTGGCTTCCCAAGCATCAACAGAAAGAG 1076
 QY 301 ValValGluIuIyGlnProTPAlaTyThrLeuSerCySglYIleValIleGlyTyHis 320
 DB 1077 GTGGTGGAG 1136
 QY 321 AsnTPGlyHisArgSerAspThrGlnAlaAsnGlyIleGlyIleGlyIleGlyIleGly 340
 DB 1137 AACCTGGGCGCATGAGAGTACAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 1196
 QY 341 ValGlyProTyraProLeuThrPleuGlyCyGlnAlaGlyPheTyraIleAspAlaGly 360
 DB 1197 GTGGGCCCTTATGTCTCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1256
 QY 361 ProProThrHisAlaPheThrProCySglYHisValCySerGlnIyIleSerAlaIleTy 380
 DB 1257 CCGGCACTCATCTTCACTCTCTTGTGACAGGTGTCTGAGAGAGTGTGCAAAATAC 1316
 QY 381 TrpSerGlnIleProLeuProHisGlyThrIleAlaPheHisAlaIleAspProPheCys 400
 DB 1317 TGGTTCAGATCCCGTGTGCTCATGAGACTCATGATTTTCAACGCTGTGCTTGTGTGT 1376
 QY 401 AlaThrGlnLeuValGlyGlnIleAsnCyIleIleIleuIlePheGlnIyProIleAsp 420
 DB 1377 GCTACACAGCTGGTGGGAGAGAAACCTGATCAAAATATTTTCCAGGTTCATTTGAC 1436

RESULT 5
 AX686405 5921 bp DNA linear PAT 29-MAR-2003
 LOCUS
 DEFINITION
 ACCESSION
 VERSION
 KEYWORDS
 SOURCE
 ORGANISM
 Homo sapiens (human)
 Homo sapiens
 Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
 Homidae; Homo.

REFERENCE
 1 Powers, S., Mu, D., Xiang, P. and Peng, Y.
 Diagnosis and treatment of cancer using mammalian pellino
 polypeptides and polynucleotides
 Patent: WO 02059611-A 3 01-AUG-2002;
 JOURNAL
 Tularik Inc. (US)
 Location/Qualifiers

source
1. .5921
/organism="Homo sapiens"
/mol_type="unassigned DNA"
/db_xref="taxon:9606"
CDS
501..1763
/note="unnamed protein product; human pellino 2"
/protein_id="CAD86576.1"
/db_xref="GI:29372144"
/translation="MPSPGSEHCAPKPEPVKGLVVLGNGALPNDRGRKRPAL
LYRPRANGVYKPSSTVHYISTPQASKASCSGHISITTSRNQTVVEYTHDQTFE
QVRSSTPIDEFVVDITSGQNDKQITOSTISRPACIVCDREPYARLPAFG
DSSKNIPLGKAKKKNPDGMDGLTNGLVMHPRGFEEQPGWREISVCGDY
TLRETSQQRGKLVSEETNVLDGSLIDLCGATLMTADGLFHTPTQGLALROE
INARPOCPVGLNTLAPPSINRKEVEKOPMAYLSGCHYGHGNNHRSQDTEANER
CPMCRTPGVPLMIGCEAGFYVDAGPPTAFTPCGHVSEKAKYWSQILPHGHHA
FHACPRCATQVLVEQNCITLITQSPID"

ORIGIN

Alignment Scores:

| Pred. No.: | 6,72e-166 | Length: | 5921 |
|------------------------|-----------|---------------|------|
| Score: | 2290.00 | Matches: | 420 |
| Percent Similarity: | 100.0% | Conservative: | 0 |
| Best Local Similarity: | 100.0% | Mismatches: | 0 |
| Query Match: | 100.0% | Indels: | 0 |
| DB: | 6 | Gaps: | 0 |

US-10-041-030-4 (1-420) x AK686405 (1-5921)

QY 1 MetPheSerProGlyGlnGluGlnHicCyAlaProAsnLysGluProValLysTYRGLY 20
DB 501 ATGTTTCCCTGGCCGAGAGAACTGGCCCCCAATAGAGAGAGCAATATACGGG 560
QY 21 GluLeuValValLeuGlyTYRAsnGlyAlaLeuProAsnGlyAspArgGlyATGARGLY 40
DB 561 GAGCTGGTGGTGGCTGGGATCAATAGTGTCTTACCATGAGATAGAGAGAGAGAAA 620
QY 41 SerArgPheAlaLeuTYRAsnGlyProLysAlaAsnGlyValLysProSerThrValHis 60
DB 621 AGTAGATTTGCCCTCAACAGCGGCCCAAGCAATAGTGTCAACCCAGCACCGTTCAT 680
QY 61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerLleSer 80
DB 681 GTGATATCCAGCGCCCGAGCATCCAGAGCTATCAGTGCAGAAAGGTCAACAGATATCC 740
QY 81 TYRThrLeuSerArgAsnGlnThrValValGlnLysTYRHisAspLysAspThrAsp 100
DB 741 TACACTTTCAGAGAAATCAGACTGTGTGTGAGATACACACTGATTAAGATACGAT 800
QY 101 MetPheGlnValGlyArgSerThrGluSerProLleAspPheValLThrAspThrLle 120
DB 801 ATGTTTCAGTGGGCGAGATCAACAGAAAGCCCTTTCATCTTGTTCACAGACACAT 860
QY 121 SerGlySerGlnAsnThrAspGlyAlaGlnLleThrGlnSerThrLleSerArgPheAla 140
DB 861 TCTGGCAGCCAGAAACGAGAGCCAGATCACACAGACACCATATCCAGGTTCGCC 920
QY 141 CysArgLleValCysAspArgAsnGluProTYRThrLArgLlePheAlaIleGlyPhe 160
DB 921 TGCAGAGTCGTGTCGACAGAGATGAACCTTACACACAGAGATTCGCCGCCGATTT 980
QY 161 AspSerSerLysAsnLlePheLeuGlyGlyLysAlaIleLysTYRAsnProAspGly 180
DB 981 GACTCTTCCAAAACATATTTCTTGGAGAAAAGCAGCAAAAGTGAAGAAACCCGACGCG 1040
QY 181 HisMetAspGlyLeuThrTrpAsnGlyValLeuValMetHisProArgGlyGlyPheThr 200
DB 1041 CACATGATGGGCTCATCTAATAGCGCTCTGTGTATGATTCACAGAGGGGCTTACC 1100
QY 201 GluGluSerGlnProGlyValLTPArgGluLleSerValCysGlyAspValTYRThrLeu 220
DB 1101 GAGGAGTCCAGCCCGGGGTCTGGCGGAGATCTCTGTCTGTGAGATGTGTACACCTTG 1160

QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyLysLeuValGluSerGluThrAsnValLeu 240
DB 1161 CGAGAAACAGGTCGGGCCGAGCAACGAGAAAGCTGTGGAAGTGAACCAACGTCCTG 1220
QY 241 GluAspGlySerLeuLleAspLeuCysGlyAlaThrLeuLeuTrpArgThrAlaAspGly 260
DB 1221 CAGAGCGGCTCCCTCAATGACCTGTGTGGGCGACCTCTCTGAGAGAACGAGATGGG 1280
QY 261 LeuPheIleThrProThrGlnLysHisLleGluAlaLeuAsnArgGlnLleAsnAla 280
DB 1281 CTTTTCATCTCCACTCGAAGACATAGAACCTCCGAGAGATTAACGCCGCC 1340
QY 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerLleAsnArgLysGlu 300
DB 1341 CGGCTCAGTGTCTGTGGGCTCAACACCTGGCTTCCCGACATCAACAGAAAGAG 1400
QY 301 ValValGluGluLysGlnProTrpAlaTYRLeuSerCysGlyHisValHisGlyTYRHis 320
DB 1401 GTGGTGAGAGAGAGAGAGCCCTGGGCAATCTCAGTGTGGCCACGTCACGGGTACAC 1460
QY 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCysArgThr 340
DB 1461 AACCTGGGCGCATCGAGTGAACAGAGGCCAAGAGAGAGTGTCCATGTGCAGAGACT 1520
QY 341 ValGlyProTYRValProLeuTrpLeuGlyCysGluAlaGlyPheTYRValAspAlaGly 360
DB 1521 GTGGGCGCCATGTGCTCTCTGTGGCTGTGGAGAGAGATTTATGTAGACGAGGA 1580
QY 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluLysSerAlaLysTYR 380
DB 1581 CCGCAACTCATCTTTCATCTCCCTGTGACAGGTGTGCTCGAGAGAGTCTGCAAAATAC 1640
QY 381 TrpSerGlnLleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
DB 1641 TGCTTCAGATCCCGTTCCTCATGGAATCATGCAATTCACCTGCTTCTCTGT 1700
QY 401 AlaThrGluLeuValGlyGlnGlnAsnCysLysLeuLlePheGlnGlyProLleAsp 420
DB 1701 GCTACAGCTGTTGGGAGAGAACTGATCAATTAATTTTCCAGGTCAATTTGAC 1760

RESULT 6
AR483568 1263 bp DNA linear PAT 14-MAY-2004
LOCUS AR483568
DEFINITION Sequence 7 from patent US 6703487.
ACCESSION AR483568
VERSION AR483568.1 GI:47246652
KEYWORDS
SOURCE Unknown.
ORGANISM Unknown.
REFERENCE 1 (bases 1 to 1263)
AUTHORS Bird,T.A. and Cosman,D.J.
TITLE Human pellino polypeptides
JOURNAL Patent: US 6703487-A 7 09-MAR-2004;
Immunex Corporation; Seattle, WA;
MOX;

FEATURES
source location/Qualifiers
1..1263
/organism="unknown"
/mol_type="genomic DNA"

ORIGIN

Alignment Scores:

| Pred. No.: | 3.94e-166 | Length: | 1263 |
|------------------------|--|---------------|------|
| Score: | 2283.00 <td>Matches:</td> <td>419 </td> | Matches: | 419 |
| Percent Similarity: | 99.8% <td>Conservative:</td> <td>0 </td> | Conservative: | 0 |
| Best Local Similarity: | 99.8% <td>Mismatches:</td> <td>1 </td> | Mismatches: | 1 |
| Query Match: | 99.7% <td>Indels:</td> <td>0 </td> | Indels: | 0 |
| DB: | 6 <td>Gaps:</td> <td>0 </td> | Gaps: | 0 |

US-10-041-030-4 (1-420) x AR483568 (1-1263)

QY 1 MetPheSerProGlyGlnGluGlnHicCyAlaProAsnLysGluProValLysTYRGLY 20

```

Db      1 ATGTTTCCCTGGCCAGAGAAACACTGGCCCTCAATAGAGAGCCAGTGAATAACGGG 60
Qy      21 GtuleuValValleuGlyYrAenglyValaleuProAenglyAaParGlyAaGly 40
Db      61 GACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
Qy      41 SerArgPheAlaLeuTyrYrAaGProLyAaAaenglyVallyAaPProSerThrValHis 60
Db      121 AGTAGATTTCCCTCTACAGAGCCGCCAGAGCAATGTGTGTGTGTGTGTGTGTGTGTGT 180
Qy      61 ValIleSerThrProGlnAlaSerIlyAaIleSerCyAlySgIyGlnHisSerIleSer 80
Db      181 GTGATATCCAGCCGCCAGGATCCAGGCTATGACGTGCAAAAGTTCACACAGATATCC 240
Qy      81 TyrThrLeuSerArgAenglyThrValValGlyIuTyrThiAaPlyAaPThrAaP 100
Db      241 TACACTTTGTCAAGAGATCAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Qy      101 MetPheGlnValGlyArgSerThrGluSerProIleAaPheValValThraaPThrIle 120
Db      301 ATGTTTCAGGTGGGACATCAACAGAAAGCCCTATGCACTTGTGTGTGTGTGTGTGTGT 360
Qy      121 SerGlySerGlnAaThrAaPGLuAaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db      361 TCTGGAGCCAGAAACAGAGCAAGCCCAATCAGACAGACAGATTCAGAGTTCCGCC 420
Qy      141 CyaaArgIleValCyaaParAaenglyuProTyrThraAaGlyIlePheAlaAaGlyPhe 160
Db      421 TGCAGAGTCGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 480
Qy      161 AaPSerSerIlyAaenIlePheLeuGlyGlyValAaIlyAaPlyAaPProAaPly 180
Db      481 GACTCTTCCAAAACATATTTCTTGAGTAAAGCAGAAAGTGAAGAAACCCCGACGGC 540
Qy      181 HiMetAaPGLyLeuThrThraaGlyValleuValMetHisPProAaGlyGlyPheThr 200
Db      541 CAATGTGATGGGCTCACTACTAATGCGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 600
Qy      201 GtugluSerGlnProGlyValTTPaRgGluIleSerValCyaaGlyAaPValTyrThrLeu 220
Db      601 GAGGAGTCCAGCCCGGGGTCTGGCGGAGATCTCTGTGTGTGTGTGTGTGTGTGTGTGT 660
Qy      221 ArgGluThrArgSerAlaGlnGlnAaGlyLyLeuValGluSerGlyThraaenValleu 240
Db      661 CGAGAAACCGGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 720
Qy      241 GlnAaPGLySerIleuIleAaPLeuCyaaGlyValAaThreuleuTTPaRgThraAaPGLy 260
Db      721 CAGAGCGGCTCCCTCATTAACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 780
Qy      261 LeuPheHisThrProThraGlnLyHisIleGluAlaLeuAaRgGlnIleAaenAlaA 280
Db      781 CTTTTCATATCTCAACTCAGAGACCATAGAACCCCTCCGGCAGAGATTAAGCGCGCC 840
Qy      281 ArgProGlnCyaaProValGlyLeuAaThrThraAaPheProSerIleAaenAaGlySgln 300
Db      841 CGGCTCAGAGTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 900
Qy      301 ValValGluGluIlyuSglnProTTPaRgIleTyrLeuSerCyaaGlyHisIlyAaGlyTyrHis 320
Db      901 GTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 960
Qy      321 AaThrPGLyHisArgSerAaPThrGluAaAaenglyuAaRgGlyCyaaPProMetCyaaRgThr 340
Db      961 AACTGGGGCCATGTGAGTACACAGAGCCCAAGAGAGAGTGTCCCATGTGTGAGACT 1020
Qy      341 ValGlyProTyrValProLeuTTPLeuGlyCyaaGlnAaGlyPheTyrValAaPGLy 360
Db      1021 GTGGGCCCCATATGTGCTCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1080
Qy      361 ProProThraHisAaPheThrProCyaaGlyHisValCyaaSerGlyuSerAlaAlyTyr 380

```

```

Db      1081 CGCGCAACTCAGCTTTCACTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1140
Qy      381 TTPSerGlnIleProLeuProHisGlyIlyThraHisAaPheHisAaIlyAaCyaaPProPheCy 400
Db      1141 TGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1200
Qy      401 AlaThrGlnLeuValGlyGluGlnAaCyaaIleAlyLeuIlePheGlnIlyProIleAaP 420
Db      1201 GCTACAGCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 1260

RESULT 7
AX298079 1263 bp DNA linear PAT 26-NOV-2001
LOCUS Sequence 7 from Patent WO0183739.
DEFINITION AX298079
ACCESSION AX298079.1 GI:17128160
VERSION AX298079.1
KEYWORDS
SOURCE Homo sapiens (human)
ORGANISM Homo sapiens
            Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini;
            Homidae; Homo.
REFERENCE 1
AUTHORS Bird, T.A. and Cosman, D.J.
TITLES Human pellino polypeptides
JOURNAL Patent: WO 0183739-A 7 08-NOV-2001;
IMMUNEX CORPORATION (US)
FEATURES
    source          1..1263
                    /organism="Homo sapiens"
                    /mol_type="unassigned DNA"
                    /db_xref="taxon:9606"

ORIGIN

Alignment Scores:
Pred. No.: 3,94e-166 Length: 1263
Score: 2283.00 Matches: 419
Percent Similarity: 99.8% Conservative: 0
Best Local Similarity: 99.8% Mismatches: 1
Query Match: 99.7% Indels: 0
DB: Gaps: 6

US-10-041-030-4 (1-420) x AX298079 (1-1263)

Qy      1 MetPheSerProGlyGlnGluGlnHisCyaaAaProAaenlyAaPProVallySlyGly 20
Db      1 ATGTTTCCCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 60
Qy      21 GtuleuValValleuGlyYrAenglyValaleuProAenglyAaParGlyAaGly 40
Db      61 GACCTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 120
Qy      41 SerArgPheAlaLeuTyrYrAaGProLyAaAaenglyVallyAaPProSerThrValHis 60
Db      121 AGTAGATTTCCCTCTACAGAGCCGCCAGAGCAATGTGTGTGTGTGTGTGTGTGTGTGT 180
Qy      61 ValIleSerThrProGlnAlaSerIlyAaIleSerCyAlySgIyGlnHisSerIleSer 80
Db      181 GTGATATCCAGCCGCCAGGATCCAGGCTATGACGTGCAAAAGTTCACACAGATATCC 240
Qy      81 TyrThrLeuSerArgAenglyThrValValGlyIuTyrThiAaPlyAaPThrAaP 100
Db      241 TACACTTTGTCAAGAGATCAGACTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGTGT 300
Qy      101 MetPheGlnValGlyArgSerThrGluSerProIleAaPheValValThraaPThrIle 120
Db      301 ATGTTTCAGGTGGGACATCAACAGAAAGCCCTATGCACTTGTGTGTGTGTGTGTGTGT 360
Qy      121 SerGlySerGlnAaThrAaPGLuAaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db      361 TCTGGAGCCAGAAACAGAGCAAGCCCAATCAGACAGACAGATTCAGAGTTCCGCC 420
Qy      141 CyaaArgIleValCyaaParAaenglyuProTyrThraAaGlyIlePheAlaAaGlyPhe 160

```


Db 421 TCGAGATCGTGGCGACAGAAATGAACCTTACACAGCAGGATATTCGCCCGCATTTT 480
Qy 161 AspSerSerIysAsnIlePheLeuGlyGluValAlaIlePheTrpIysAsnProAspGly 180
Db 481 GACTCTTCCAAAACATATTTCTTGGAGTAAAGGACCAAAAGTGAAGAAACCCGACGGC 540
Qy 181 HisMetAspGlyLeuThrTrpAsnGlyValIleValMetHisProArgGlyGlyPheThr 200
Db 541 CACATGATGGGCTCACTACTAATGGCGTCTGGTGTATGATCCACAGAGGGGGCTTACC 600
Qy 201 GluGluSerGlnProGlyValITrpaGluIleSerValCysGlyAspValIlyThrLeu 220
Db 601 GAGAGAGCCAGCCCGGGGCTGGCGCGAGATCTCTCTGTGGAGATGGTACACCTTG 660
Qy 221 ArgGluThrArgSerAlaGlnIleAspLeuGlyValIleValGluSerGlnThrAsnValLeu 240
Db 661 CGAGAAACAGGCTCGGCCACGACGAAAGAGAAAGCTGGTGAAGATGAGACCAAGCTCTG 720
Qy 241 GlnAspGlySerLeuIleAspLeuGlyValIleThrLeuLeuTrpArgThrAlaAspGly 260
Db 721 CAGGAGCGCTCCCTCATTTGACCTGTGTGGGGCCACTCTCCCTGGAGAACAGAGAGG 780
Qy 261 LeuPheHisThrProThrGlnIlyHisIleGluAlaLeuArgGlnIleIleAsnAla 280
Db 781 CTTTTCATCTCCAACTCAGAGACATAGAAAGCCCTCCGCGAGATTAACGCCGCC 840
Qy 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgGlyGlu 300
Db 841 CGGCTCAGTGTCTGTGGGCTCAACACCTTGCTCCCGACATCAACAGAAAGAG 900
Qy 301 ValValGluGluIlyGlnProTrpAlaIlyLeuSerCysGlyHisValIleGlyThrHis 320
Db 901 GTGGTGAAGAGAAAGAGCCCTGGGCAATCTCATGTTGTGGCCACGTGACGGGTACAC 960
Qy 321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGluArgGlyCysProMetCysArgThr 340
Db 961 AACTGGGGCCATCGAGTGAACGAGGCAACGAGGAGAGTGTCCCATGTGCGAGACT 1020
Qy 341 ValGlyProTrpValProLeuTrpLeuGlyCysGluAlaGlyPheTrpValAspAlaGly 360
Db 1021 GTGGGCCCCATGTGCTCTGTGGCTGTGGCTGTGAGGACAGATTTATGTAGACGGAGA 1080
Qy 361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluIlySerAlaIlyThr 380
Db 1081 CCGGCAACTCATGCTTCACTCCCTGTGACAGCTGTGCTGGAGAGTCTGCAAAATAC 1140
Qy 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400
Db 1141 TGGTCTCAGATCCGTTGCTCATGAACTCATTCATTCACGCTGCTGCCCTTCTGT 1200
Qy 401 AlaThrGlnLeuValGlyGluIleAsnCysAlleIleLeuPheHisGlnIlyProIleAsp 420
Db 1201 GCTACACAGCTGTGGGAGCAAACTGATCAATTAATTTCCAAAGTCCCAATTGAC 1260

RESULT 8
AR483567 1260 bp DNA linear PAT 14-MAY-2004
LOCUS AR483567
DEFINITION Sequence 5 from patent US 6703487.
ACCESSION AR483567
VERSION AR483567.1 GI:47246651
KEYWORDS
SOURCE Unknown.
ORGANISM Unclassified.
REFERENCE 1 (bases 1 to 1260)
AUTHORS Bird, T.A. and Cosman, D.J.
TITLE Humanellino polypeptides
JOURNAL Patent: US 6703487-A 5 09-MAR-2004;
Immunex Corporation; Seattle, WA;
MOX;
FEATURES
source Location/Qualifiers
1..1260

ORIGIN /organism="unknown"
/mol_type="genomic DNA"
Alignment Scores:
Pred. No.: 7,216-158 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.4% Conservative: 10
Best Local Similarity: 95.0% Mismatches: 10
Query Match: 95.0% Indels: 1
DB: 6 Gaps: 1
US-10-041-030-4 (1-420) x AR483567 (1-1260)
Qy 1 MetPheSerProGlyGlnGluGlnHisCysAlaProAsnIlyGluProValIlyTrpGly 20
Db 1 ATGTTTTCCTCCGGCCAGAGAGAACCCACGCCCCCAACAGAGAGCGGTGAATTCGGG 60
Qy 21 GluLeuValValIleGlyIlyTrpAsnGlyAlaLeuProAsnGlyAspArgGlyArgArgIys 40
Db 61 GACCTGTGGTCTGGGGTAAATGATGCTTAACTTAATGTATGACAGGGGCGAGAGAAA 120
Qy 41 SerArgPheAlaLeuTrpIysArgProIyAlaAsnGlyValIlyPheProSerThrValHis 60
Db 121 AGCAGATTGGCTCTTAAGCGGACCTACGCGCATGGTGTCAAAACCCAGCAAAATCCAC 180
Qy 61 ValIleSerThrProGlnAlaSerIlyAlaIleSerCysIlyGlnHisSerIleSer 80
Db 181 ATGCTTCCACACCAACAGCGCTCCAGGCCATCTCCAGAGACATCACAGATATG 240
Qy 81 TyrThrLeuSerArgAsnGlnThrValValIleGlyIlyThrHisAspIlyAspThrAsp 100
Db 241 TACACGTTGTCAAGAGCCAGACGATAGTGTGATGATCAACACATTAAGACAGGAC 300
Qy 101 MetPheGlnValGlyArgSerThrGluSerProIleAspPheValIlyThrAspThrIle 120
Db 301 ATGTTTCAGGTGGAGGTCAACAGAAAGCCCATTTGATCTGTGTCAACAGCGGT 360
Qy 121 SerGlySerGlnAsnThrAspGlnAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
Db 361 TCCGGCGCTCAAGAACAAAGT---GCCAGATCACAGACCACTCTCTAGGTTCCGA 417
Qy 141 CysArgIleValCysAspArgAsnGluProIlyThrAlaArgIlePheAlaIleGlyPhe 160
Db 418 TGCAGAGTGTGTGTACAGAAAGACCATTAACAGACGATATTCGCGGACGATTC 477
Qy 161 AspSerSerIysAsnIlePheLeuGlyGluValAlaIlePheTrpIysAsnProAspGly 180
Db 478 GATTTCTCCAAAATATTTCTTGTGAGAGAAAGCAAAATGAGAAAACCTGATGGA 537
Qy 181 HisMetAspGlyLeuThrTrpAsnGlyValIleValMetHisProArgGlyGlyPheThr 200
Db 538 CACATGATGAGATCTCACTACATGATGCTCTTAATGATGACCCCGAGAGAGCTTCACC 597
Qy 201 GluGluSerGlnProGlyValITrpaGluIleSerValCysGlyAspValIlyThrLeu 220
Db 598 GAGGAATCCAGCCTGAGAGTCTGAGAGAGATCTCTGTGTGGAGATGTACACTTG 657
Qy 221 ArgGlnThrArgSerAlaGlnIleAspGlyIlyValIleValGluSerGlnThrAsnValLeu 240
Db 658 CGAGAGACAGGCTCGGCCACGAGAGGAGAAAGCTGTGGAAGTGAAGACCAAGCTCTG 717
Qy 241 GlnAspGlySerLeuIleAspLeuGlyValIleThrLeuLeuTrpArgThrAlaAspGly 260
Db 718 CAGAGCGCTCCCTCATTTGACCTGTGTGGGGCCACTCTCTCTGGAGAACCGGAGATGGC 777
Qy 261 LeuPheHisThrProThrGlnIlyHisIleGluAlaLeuArgGlnIleIleAsnAla 280
Db 778 CTTTTCACGCTCTCACTCAGAAAGCATAGAAAGCCCTCCGAGAGAGATCAATGACAGC 837
Qy 281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgGlyGlu 300
Db 838 CGACCCAGTGCCTGGGGCTTAAACACCTGTGCTTCCCGCATCAACCGGAAGGAA 897

| | | | | |
|----|--|------|---|------|
| Oy | | 301 | AAlValGlugluYwylvgInPrOTrThLaTyrrLeusSerCySglVhlHlGlyThyrtIs | 320 |
| Dd | | 898 | GfGGtGAAGAGAAAGCACCCCTGGGCATACCTTGAeTCGGCCAGTGCAAGGCTAACAC | 957 |
| Oy | | 321 | ASnTrpgIyhISArGSerApThrCluAaSnGIuATrGluCYSPromECysArGrThr | 340 |
| Dd | | 958 | AGCGGGGGCCATCGGAGGAGCGCGGAAACCAGAGAGGAGGTGCCATGTGCAGACT | 1017 |
| Oy | | 341 | VAlGlgYProlYrValProleuTrpleuGLYCYSgUwlaaglyPheTYrValAspAlagly | 360 |
| Dd | | 1018 | GTGGGGCCCCTAGTCCTCCTCTGTGGCTGTGTGAGGAGATTATTAAGTCATCGCGGA | 1077 |
| Oy | | 361 | PRoPrOThrhISAlAPHeThPrOCYsAgIyhISVAlCYSeSerGluVSerAlaVetTyR | 380 |
| Dd | | 1078 | CCCCCACTCACGCTTTACACCCCCTCGGGGACGCTGTTCAGANAAGTCGCAAGTAC | 1137 |
| Oy | | 381 | TpSGerglnIEProleuProHlaGlyThrhISAlAPHeHISAlaAlaCYSProBheCYs | 400 |
| Dd | | 1138 | TGTGTGCGAGATCCCACTGCCCCACGGAAGCAGCGGTTTCATGCGCGCTGTCCGTTCG | 1197 |
| Oy | | 401 | AlAtHrGlnLeuValIGLyGLInAnScYelleyLysleuIIepneGlnGlyProIlEaSp | 420 |
| Dd | | 1198 | GCCAAGCAGCTGTGTGTGGAAACAAGATCATCAATTAATTTCAGAGTCCAGTGCAC | 1257 |

RESULT_9
AX298077 LOCUS AX298077 1260 bp DNA linear PAT 26-NOV-2001

DEFINITION Sequence 5 from Patent WO0183739.

ACCESSION AX298077

VERSION AX298077.1 GI:17128159

KEYWORDS Mus musculus (house mouse)

SOURCE ORGANISM Mus musculus

REFERENCE AUTHORS Buxarjoca, Metazoa; Chordata; Craniata; Vertebrate; Euteleostomi;
TITLE Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
JOURNAL Sclurognathi; Muridea; Muridae; Murinae; Mus.
IMMUNEX CORPORATION (US) Patented: NO 0183739-A 5 08-NOV-2001;

FEATURES location/Qualifiers
source 1..1260 /organism="Mus musculus"
/mol_type="unassigned DNA"
/db_xref="taxon:10090"

ORIGIN Alignment Scores:
Pred. No.: 7,21e-158 Length: 1260
Score: 2175.50 Matches: 399
Percent Similarity: 97.4% Conservative: 10
Best Local Similarity: 95.0% Mismatches: 10
Query Match: 95.0% Indels: 1
DB: Gaps: 1

US-10-041-030-4 (1-420) x AX298077 (1-1260)
Oy 1 MetPheSerPrOGlyGInGluGluHiSCySaIAProAenLYBgUProvAllySTryGly 20
Dd 1 ATGTTTTCCCGGGCCAGAGAGAACCCAGCCGCCCAACAAGAGGCCGTGAATAACGGG 60
Oy 21 GlueuVeAlIVallEugLYTyRAnglYAlaleuPrOAengIYAaPaRGlyYARayrlys 40
Dd 61 GAGCTGGTGTCTCGGGGTACATGTCTTACTTAATGTGACAGAGGGGAGAGGAAA 120
Oy 41 SerArgPheaAlaEuTYrLYsArgPrObYsaIAaengIYAllyPSerThrVAHis 60
Dd 121 AGCAGATTTCCTCTATTAAGGGGACCTACGCCCACTGTGGTGTCAAACCAGACCAATCCAC 180
Oy 61 ValIleSerThrPrOGlnAlaSerIyaAlaIleSerCYblyvGlyGlnHISerIleSer 80
::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: :::: ::::

| | | | |
|----|------|--|------|
| Db | 181 | ATGCTTCCACACACAGAGGCTCCAAAGGCATCAAGCTCCAGAGACATCAAGCATATCG | 240 |
| Qy | 81 | TYRTHLEUSERARGAENGINTHRVALVALVALGINTYRTHIASAPLYASPTHRASP | 100 |
| Db | 241 | TACACGTTGTACACGAGACGACGCGTAGTGTGAGTACACACAGTATTAAGACACGGAC | 300 |
| Qy | 101 | METHEGINVALIGLYASGETHCHUSERPROILEAPHPHEVALVALTHASPTHRIE | 120 |
| Db | 301 | ATGTTTCAGCGGGCAGGTCAACGAAAGCCCATTAATTCCTCGGGTCACAGACCGGTT | 360 |
| Qy | 121 | SERGLYSERGINAENTHRASPGVALAGNILETHRGINSERTHRILESERARGPHEA | 140 |
| Db | 361 | TCCGGCGGTCAGAAACGAAGAT--GCCAGATCCACAGAGCACATCTTAGGTTGCCA | 417 |
| Qy | 141 | CYAAAGLILEVALCYAASPARASANGINPROTYRTHRILAARGILEPHEVALAIGLYPHE | 160 |
| Db | 418 | TGCAGAGATCGTGCTGACAGGAGACGACCATATACAGACGCATATTCGCGGACGATTC | 477 |
| Qy | 161 | ASPSERSELYSAANILEPHELEUGLYGULVALAALAYSTRPLYSAENPROASPGLY | 180 |
| Db | 478 | GATTCCTCCAAAATAATCTTTCTTGAGAGAGAAACACCAATAATGAAAMACCTCGATGGA | 537 |
| Qy | 181 | HISWEASPGLYLEUTHRTHRASNGLYVALILEUVALIMETHISPROARGGLIPLYPETHR | 200 |
| Db | 538 | CACATGATGGACTCACTACCAATAGTGCTCTAGTAGTACACCCGCAAGAGGCTTCAAC | 597 |
| Qy | 201 | GLUGLUSERGINPROGLIYVALTRPARAGLUILESERVALCYSGIYASPVALTyrTHLEU | 220 |
| Db | 598 | GAGGAATCCAGCGCTGAGGTCTGAGAGAGATCTCTGTGTGGGAGATGTGTACACTTG | 657 |
| Qy | 221 | ARGGLUTHRARGSERVALAGINGLNAARGLYLYLEUVALIGLUSERGLUTHRASVALLEU | 240 |
| Db | 658 | CGAAGACACAGGTGGGCCACAGCAGAGGGGAAAGCTGTGTGGAATGTGAGCAACAGTCTGTG | 717 |
| Qy | 241 | GLINASPGLYSERLEULIAPSPLEUCYSGLYVALATHRLEULEUTHRATGTHRALASPGLY | 260 |
| Db | 718 | CMAAGCGCTCCCTCATATTACCTGTGTGGGCCACTCTCTCTGAGAACCGCAGATGAC | 777 |
| Qy | 261 | LEUPHEHISTHRPROTHRGINLYSHSILEGLUALALEUARGGLINUILEASNALA | 280 |
| Db | 778 | CTTTTTCACGCTCTCACTCAGAAACACATAGAACCCCTCGGCGAGGATCAATGACGCC | 837 |
| Qy | 281 | ARGPROGLINCYSPROVALGYLEUASHTHRLAALAPHEPROSERILEASNALYGLIU | 300 |
| Db | 838 | CGAACCCAGTGGCCCCGTGGGCTTTAAACCTCTGCTCCACATCAACCGAAGGAA | 897 |
| Qy | 301 | VALVALIGLUGIUYWSGINPROTPALALETYRLEUSERCYSGIYHISVALHISGLYTHIS | 320 |
| Db | 898 | GTGGTGTGAAGAAAGACAGCCCTGGGCAATCTGAGCTGGCCCATGTGCAGCGGCTAACAC | 957 |
| Qy | 321 | ASNTTRPGLYHISARGSERAPSTHGLUALASANGIUAARGIUCYSPROMETCYASRGTTHR | 340 |
| Db | 958 | AGCTGGGGCCATTCGAGACGACCGGAGACAGACAGAGGAGTGTCCCATGTGCAGGACT | 1017 |
| Qy | 341 | VALGLYPROTYRVALPROLEUTRPLEUGLYCYSGIUALAGLYPHETYRVALASPALAGLY | 360 |
| Db | 1018 | GTGGGCCCCCTAAGTCCCTCTCTGTGGCTGGGCGTGTAGGACAGGATTTTATGTGTCAGCGGGA | 1077 |
| Qy | 361 | PROPTHRTNHSIALAPETHRPROCYGLYHISVALCYSSERGIUYSSERVALYSTR | 380 |
| Db | 1078 | CCCCCACTACCGGCTTTCACCCCCTGGGGGACGTCGTGTTCAGAAAGTCTGGCAAGTAC | 1137 |
| Qy | 381 | TRPSERGINILEPROLEUPROHISGLYTHRTHISALAPHEHISALAAIACYSPROPHCY | 400 |
| Db | 1138 | TGCTTCGAGATCCACTGCCCCACAGAACGACCGTTTATGCGGCTGTCCGTTCTGC | 1197 |
| Qy | 401 | ALATHRGINLEUVALIGLYGLUGINASRCYSILEYALEUILEPHEGLINGLYPROILEASP | 420 |
| Db | 1198 | GCCACGACGAGCTGTGTGTGAAACAGACATGCATCAATATGATTTTCCAAAGGTCCAGTGAC | 1257 |

DEFINITION Xenopus laevis MGC80329 protein, mRNA (cDNA clone MGC:80329
IMAGE:5074048), complete cds.
ACCESSION BC072891
VERSION BC072891.1 GI:49119163
KEYWORDS MGC.
SOURCE Xenopus laevis (African clawed frog)
ORGANISM Xenopus laevis
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
Xenopodinae; Xenopus; Xenopus.
REFERENCE 1 (bases 1 to 1646)
Klein, S.L., Strausberg, R.L., Wagner, L., Pontius, J., Ciffton, S.W.
and Richardson, P.
Genetic and genomic tools for Xenopus research: The NIH Xenopus
Initiative
JOURNAL Dev. Dyn. 225 (4), 384-391 (2002)
PUBMED 12454917
REFERENCE 2 (bases 1 to 1646)
Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, P.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Heide, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Shapleton, M., Soares, M.B., Bonaldo, M.F., Cabavant, T.L.,
Scheetz, T.E., Brownstein, M.J., Usdin, T.B., Toshiyuki, S.,
Chenrai, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Abrams, R.D., Mulhaly, S.J., Bosak, S.A., McGowan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hultk, S.W.,
Villalón, D.R., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahney, J., Halton, E., Kettner, M., Madan, A., Rodriguez, S.,
Sanchez, A., Whitting, M., Madan, A., Young, A.C., Rodriguez, Y.,
Bouffard, G.G., Blakesley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Smalins, D.E.,
Scherer, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16699-16903 (2002)
JOURNAL 12477932
PUBMED
REFERENCE 3 (bases 1 to 1646)
Klein, S. and Gerhard, D.S.
Direct Submission
Submitted (07-JUN-2004) National Institutes of Health, Xenopus Gene
Collection (XGC), National Institute of Child Health and Human
Development, 6100 Executive Boulevard, Room 4B01, Rockville, MD
20892-7510, USA
REMARK
COMMENT NIH-MGC Project
Contact: XGC help desk
Email: cgabs-remail.nih.gov
Tissue Procurement: Dr. Igor David
cDNA Library Preparation: Life Technologies, Inc.
cDNA Library Arrayed by: The I.M.A.G.E. Consortium (LNL)
DNA Sequencing by: Genome Sequence Centre,
BC Cancer Agency, Vancouver, BC, Canada
info@cgsc.bc.ca
Steve Jones, Sarah Barber, Mabel Brown-John, Yaron Butterfield,
Andy Chan, Steve S. Chand, William Chow, Alison Cloutier, Ruth
Featherstone, Malachi Griffith, Ohi Griffith, Ran Guin, Nancy Liao,
Kim MacDonald, Amara Masson, Mike R. Mayo, Josh Moran, Ryan Morin,
Teika Olson, Diana Palmquist, Anca Petrescu, Anna Lissa Prihbu,
Parvaneh Saeedi, Jr Santos, Angeliue Scherch, Ursula Skalska,
Duane Smilun, Jeff Scott, Miranda Tsai, George Yang, Jacquiue
Schein, Asim Siddiqui, Rob Holt, Marco Marra.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: <http://image.llnl.gov>
Series: IRAX Plate: 152 Row: h Column: 6
This clone was selected for full length sequencing because it
passed the following selection criteria: Hexamer frequency ORF
analysis, Similarity but not identity to protein.
location/Qualifiers
1..1646

ORIGIN
Alignment Scores:
Pred. No.: 6,84e-155 Length: 1646
Score: 2138.50 Matches: 383
Percent Similarity: 96.9% Conserved: 24
Best Local Similarity: 91.2% Mismatches: 12
Query Match: 93.4% Indels: 1
DB: 5 Gaps: 1
US-10-041-030-4 (1-420) x BC072891 (1-1646)
QY 1 MetPheSerProGlyngIngluGluHsiCyBaLaProAsnLysGluProVallyrTygLy 20
DB |||||CCCCCGAGTCAAGGAAACATGGGCCCCGAGGAGCCCGTGAATACGGA 395
QY 21 GluLeuValValleuGlyTyraGngLyAlaLeuProAsnGlyYAspArgLyATGATG 40
DB |||||GAGCTGTGTGTGGGGTACACAGGCTGTCTCCAAATGGGATGAGGAAGCCGGA 455
QY 41 SerArgPheAlaLeuTyTyryATGATGProLYaLaAsnGlyYValYAspSerThrValHis 60
DB 456 AGCAGATTTCACCTCTTCAACCCGCCCAATCTMAAGGGGTGAGACCCAGCTGTTCAT 515
QY 61 ValIleSerThrProGlnAlaSerLyBaLaIleSerCyLeuagLyGlnHisSerIleSer 80
DB 516 GTGATTTCAACCCCGAGGCTTCCAGGCAATGACAGCAGAGCCCAACACACATCTCA 575
QY 81 TyrThrLeuSerArgAsnGlnThrValValaGluTyThrHisAspYAspThrAsp 100
DB 576 TTCACATTATCCAGAGAGTACAGCTGTGTGTGGAATACCCATGACAGACAGACAGAT 635
QY 101 MetPheGlnValaGlyArgSerThrGluSerProIleAspPheValValThrAspThrIle 120
DB 636 ATCTTTTCAGATTGTGAAGGTCAACAGAAAGCCCATTTGATGTAATCTGTATACGGT 695
QY 121 SerGlySerGlnAsnThrAspGluValaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 696 TCTGGCAACGAGAAC--GATGAACACAGATACCAAAATTCAAATTCAGATTGCA 752
QY 141 CysArgIleValCyAspArgAsnGluProTyThrAlaArgIlePheAlaIleGlyPhe 160
DB 753 TGTAGATATGTCGCAATGAGAACCCGCGTACAGCTGATATTTGTGCGAGATTT 812
QY 161 AspSerSerIleAsnIlePheLeuGlyGlyuLeuAlaAlaLysTrpYAspAspArgGly 180
DB 813 GACTCCTCCAAATATATTTCTTGTGTGAGAAAGCGCCAAAGTGAAGAAATCTGACGGT 872

Alignment Scores:

Pred. No.: 4,05e-154 Length: 1936
 Score: 2129.50 Matches: 384
 Percent Similarity: 96.7% Conservative: 22
 Best Local Similarity: 91.4% Mismatches: 13
 Query Match: 93.0% Indels: 1
 Gaps: 1

US-10-041-030-4 (1-420) x BC063200 (1-1936)

```

QY 1 MetPheSerProGlyngluGluHisCyvalaProbaNlysgluProVallyrGly 20
DB 365 ATGTTTCCCCCGAGTCAGAGAGACAGTGGCCCCCGAGTAAGAGCCCGCAATATAGCA 424
QY 21 GluLeuValIleuGlyTyrAenGlyValaLeuProbaNlysgluPargyls 40
DB 425 GAGCTGGTGGTGGGGGTACACGGCTGTCTCCCAATGGGGATCGAGAAAGCGGAAA 484
QY 41 SerArgPheAlaLeuTyrIlysaRProLyalaAsnGlyVallyrProSerThrValHis 60
DB 485 AGCAGATTTCACCTTTAAAGCCCTTAATCCATGGGGTGAAGCCAGCAGTTCAT 544
QY 61 ValIleSerThrProGlnAlaSerLyalaIleSerCyalygIlnHisSerIleSer 80
DB 545 GTGATTTCAACCCACAGGCTTCCAGAGCAATCGACAGAGGCCAAACAGCATCTCA 604
QY 81 TyrThrIleuSerArgaenGlnThrValValIleGluTyrThrHisaRplysaRThrAsp 100
DB 605 TTCCATATATCCCAAAAGTCAAGACTGGTGTGGTAATACCCCAATGACAAAGACACAGAT 664
QY 101 MetPheGlnValGlyArgSerThrGlySerProIleAspPheValValThrAspThrIle 120
DB 665 ATGTTTCAAGTGGAAAGTCAACGAAGTCCATTACTTTGATGACTGACACGGTT 724
QY 121 SerGlySerGlnaenThrAspGluValaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB 725 TCTGGCAACACGAAT---GATGAACAACAGATTAACAAGTCAATCTCAGATTGGC 781
QY 141 CyabArgIleValCyabaRArgaenGluProTyrThrIlaArgIlePheAlaIleGlyPhe 160
DB 782 TGCAGATATCTGTGACAGAACCTCTCCGTAACCCCTGAATATTTGTCGAGATTT 841
QY 161 AspSerSerLyalaenIlePheLeuGlyValIleValaIleAspTPLYsaRProAspGly 180
DB 842 GACTCCCTCCAAAATATATTTCTTGGTGAAGAACGGCCCAAGTGAAGATCCGATGCT 901
QY 181 HisMetAspGlyLeuThrThrAsnGlyValIleuValMetHisProArgIlyGlyPheThr 200
DB 902 CATATGATGAGACTAACCTAACATAGTGTCTTGTGATGATCCGAAGGAGGGGTTACT 961
QY 201 GluGluSerGlnProGlyValITPArgGluIleSerValCyabGlyAspValTyrThrLeu 220
DB 962 GAAGAAATCCAGGCTGAGATAGCGGGAATTCAGTCTGTGAGATGTATATATCACTTA 1021
QY 221 ArgGluThrArgSerAlaGlnGlnArgGlyIleLeuValGluSerGluThrAsnValLeu 240
DB 1022 AGAGAAACCGATTTGCTCAGCAAGAGGCAATCTTGGAATAATGAACCAATATCTG 1081
QY 241 GlnAspGlySerLeuIleAspLeuCyabGlyValaThrLeuLeuTyrArgThrAlaAspGly 260
DB 1082 CAGATGTTCTCTGTGACTTATGTGGCCACCTCTTTGGCAACGACGATGCT 1141
QY 261 LeuPheHisThrProThrGlnLysHisIleGluAlaLeuArgGlnGluIleAsnAla 280
DB 1142 TTGCTCATACACGACTCAAAAACATTTGAGCTTGAACGGAATTAATAGCTGCC 1201
QY 281 ArgProGlnCysProValaGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLyGlu 300
DB 1202 AGGCTCAGAGCCCTGTGGATTAACAATTACCTTTCCAGATTAACAACGGAAAGAT 1261
QY 301 ValValGluGluLysGlnProTPAlaTyrLeuSerCyabGlyHisValHisGlyTyrHis 320
  
```

```

DB 1262 GTCTAGAAAGAAAGACAGCTTGAGCTTACCTGACCTGTGGTATGTCAGCGCTATCAT 1321
QY 321 AsnTyrGlyHisArgSerAspThrGluAlaAsnGluArgGluCysProMetCyabArgThr 340
DB 1322 AACCTGGGTATGTCAGTGAACAGAAAGTCAATGAACGGACAGTGTCCATGTGTAAACA 1381
QY 341 ValGlyProTyrValProLeuThrPleuGlyCyabGluAlaGlyPheTyrValaAspAlaGly 360
DB 1382 GTTGGTCCCTATGCTCTATGCTGTGGTGTGAAGACAGATTTCAAGTATGCTGGA 1441
QY 361 ProProThrHisAlaPheThrProCyabGlyHisValaCyabSerGluLysSerAlaLeuTyr 380
DB 1442 CCTCTTACTCACCTTTCAAACTTGTGACAGCTCAGCTCAGAAAAATCTGCAATAT 1501
QY 381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaIlaCysProPheCys 400
DB 1502 TGGTTCAGATTCACCTGCCACATGGTACACAGCGCTTTCATCTGCTTCTTCTGC 1561
QY 401 AlaThrGluLeuValaGluGlnaenCyabIleLeuLeuIlePheGlnGlyProIleAsp 420
DB 1562 GCAATTAATTAATGACTGGAAGAAAGAGCTGTGCAAACTTATTTCCAGGGCCCTGTGAC 1621
  
```

RESULT 12
 BD190327
 LOCUS BD190327 1717 bp DNA linear PAT 17-JUL-2003
 DEFINITION BD190327 1 GI:33000066
 ACCESSION BD190327
 VERSION BD190327.1 GI:33000066
 KEYWORDS WO 03008589-A/67
 SOURCE Mus musculus (house mouse)
 ORGANISM Mus musculus

Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia;
 Sciurognathi; Muridae; Murinae; Mus.
 1 (bases 1 to 1717)
 AUTHORS Matsuzaki, O., Matsuda, A., Nagano, Y. and Suzuki, N.
 TITLE Btk1 phosphorylation related gene
 JOURNAL Patent: WO 03008589-A 67 30-JAN-2003;
 ASahi Kasei Corp, OSAWU MATSUZAKI, AKIO MATSUDA, YUKIKO NAGANO, NAOMI SUZUKI

COMMENT OS Mus musculus (mouse)
 PN WO 03008589-A/67
 PD 30-JAN-2003
 PE 15-JUL-2002 WO 2002JP007174
 PR 18-JUL-2001 JP 01P 218204, 31-AUG-2001 JP 01P 263450 PR
 21-JAN-2002 JP 02P 012176
 PI OSAWU MATSUZAKI, AKIO MATSUDA, YUKIKO NAGANO, NAOMI SUZUKI PC
 C12N15/54, C12N9/12, G01N33/50, G01N33/15, C07K16/40, A61K39/385, PC
 A61K31/711
 CC Btk1 phosphorylation related gene
 FH Key location/Qualifiers
 FT CDS (218) . (1474) .

FEATURES
 source location/Qualifiers
 1..1717
 /organism="Mus musculus"
 /mol_type="genomic DNA"
 /db_xref="taxon:10090"

ORIGIN

Alignment Scores:

Pred. No.: 8.55e-153 Length: 1717
 Score: 2111.50 Matches: 388
 Percent Similarity: 95.2% Conservative: 12
 Best Local Similarity: 92.4% Mismatches: 19
 Query Match: 92.2% Indels: 1
 Gaps: 1

US-10-041-030-4 (1-420) x BD190327 (1-1717)

```

QY 1 MetPheSerProGlyngluGluHisCyvalaProbaNlysgluProVallyrGly 20
DB 218 ATGTTTCCCCCGAGTCAGAGAGACAGTGGCCCCCGAGTAAGAGCCCGGTAATATAGCA 277
  
```

```

QY 21 G|U|L|E|U|V|A|L|E|U|G|L|Y|T|Y|R|A|N|G|L|Y|A|L|E|U|P|R|O|A|N|G|L|Y|A|S|P|R|G|L|Y|R|G|L|Y|S 40
DB 278 G|A|C|T|G|T|G|T|C|T|G|G|G|T|C|A|A|T|G|T|G|T|T|A|C|T|T|A|T|G|T|A|C|A|G|G|G|C|A|A|G|A|A 337
QY 41 S|E|R|A|R|P|H|A|L|A|L|E|U|T|Y|R|L|Y|R|A|R|P|R|O|L|Y|S|A|A|S|N|G|L|Y|A|L|Y|S|P|R|O|S|E|R|T|H|A|L|S 60
DB 338 A|G|A|G|A|T|T|A|A|C|C|T|T|A|T|A|G|G|G|A|C|C|T|A|G|C|A|G|G|T|G|T|C|A|A|C|C|C|A|C|A|A|T|C|C|A|C 397
QY 61 V|A|L|I|S|E|R|T|H|P|R|O|G|L|A|S|E|R|L|Y|S|A|A|L|I|S|E|R|C|Y|L|S|G|L|Y|G|L|N|H|I|S|E|R|I|L|E|S|R 80
DB 398 A|T|G|T|T|C|C|A|C|A|C|A|G|G|G|C|C|A|G|G|C|C|A|G|C|C|A|G|C|C|C|A|G|A|G|A|C|A|C|A|T|A|T|C|G 457
QY 81 T|Y|R|T|H|L|E|U|S|E|R|A|R|G|A|R|G|I|N|T|H|A|L|V|A|L|G|U|T|Y|R|H|I|S|A|P|R|Y|S|A|R|P|R|A|R|P 100
DB 458 T|A|C|A|C|G|T|T|C|A|C|A|G|G|C|C|A|G|G|T|A|G|T|G|T|G|A|G|T|A|C|A|C|A|C|A|T|A|A|A|G|C|A|C|A|G|A|C 517
QY 101 M|E|T|P|H|E|G|L|N|A|L|G|L|Y|R|S|E|R|T|H|G|L|U|S|E|R|P|R|O|I|L|E|S|R|P|H|E|V|A|L|T|H|R|A|R|P|R|I|L|E 120
DB 518 A|T|G|T|T|C|A|G|T|G|G|G|A|G|T|C|A|C|A|G|A|A|G|C|C|A|T|T|A|C|T|T|G|T|G|T|C|A|C|A|G|A|C|G|G|T 577
QY 121 S|E|R|G|L|Y|R|S|E|R|G|L|N|A|R|H|A|R|P|G|L|U|A|G|L|N|I|L|E|R|T|H|G|L|N|S|E|R|T|H|I|S|E|R|A|R|P|H|E|A|L 140
DB 578 T|C|C|G|G|G|G|T|C|A|G|A|C|G|A|A|G|A|T|---G|C|C|A|G|A|T|C|A|C|A|G|A|C|A|C|A|T|C|T|A|G|T|T|C|G|C|A 634
QY 141 C|Y|E|R|I|L|E|V|A|L|C|Y|S|A|R|P|A|R|G|A|R|G|I|N|P|R|O|T|Y|R|T|H|A|R|G|L|E|P|H|E|A|L|A|G|L|Y|P|H|E 160
DB 635 T|G|C|A|G|A|T|G|T|G|T|G|T|G|A|C|A|G|A|A|C|G|A|C|C|A|T|A|C|A|G|C|A|T|A|T|T|C|G|G|G|C|A|G|A|T|T|C 694
QY 161 A|S|P|R|S|E|R|Y|S|A|R|I|L|E|P|H|E|U|G|L|Y|L|U|Y|S|A|A|L|A|L|Y|R|P|L|Y|S|A|R|P|R|O|A|R|P|G|L|Y 180
DB 695 G|A|T|T|T|C|C|A|A|A|A|A|T|C|T|T|T|T|G|A|G|A|A|A|G|C|A|A|A|A|A|G|A|A|A|A|C|C|T|G|A|T|G|A 754
QY 181 H|I|S|E|R|A|R|P|G|L|Y|L|E|U|T|H|R|A|R|G|L|Y|V|A|L|E|U|V|A|L|E|U|T|H|I|S|P|R|O|R|G|L|Y|G|L|Y|P|H|E|R|T|H 200
DB 755 C|A|C|A|T|G|A|T|G|A|C|T|A|C|T|A|C|A|C|G|G|G|T|C|T|G|T|G|A|T|G|C|A|C|C|G|C|A|G|A|G|G|C|T|T|C|A|C|C 814
QY 201 G|L|U|G|L|S|E|R|G|L|N|P|R|O|G|L|Y|A|L|T|P|R|A|R|G|L|U|I|L|E|S|R|V|A|L|C|Y|S|G|L|Y|R|A|R|P|V|A|L|Y|R|T|H|L|E|U 220
DB 815 G|A|G|A|A|T|C|C|C|A|G|G|C|T|G|A|G|T|G|A|G|A|G|A|T|C|T|G|T|G|T|G|G|G|A|G|T|A|T|A|C|A|C|T|T|G 874
QY 221 A|R|G|L|U|T|H|R|A|R|S|E|R|A|R|G|I|N|A|R|G|L|Y|L|U|Y|S|A|L|E|U|V|A|L|G|L|S|E|R|G|L|U|T|H|R|A|R|P|V|A|L|L|E|U 240
DB 875 C|G|A|G|A|A|C|C|A|G|T|C|G|G|C|C|A|G|C|A|G|G|G|G|A|A|G|T|G|T|G|A|A|G|T|G|A|C|A|C|A|C|G|T|C|T|G 934
QY 241 G|L|N|A|R|P|G|L|Y|R|S|E|R|L|E|A|R|P|L|E|U|C|Y|S|G|L|Y|A|L|T|H|L|E|U|T|P|R|A|R|G|L|U|A|R|P|G|L|Y 260
DB 935 C|A|A|G|A|C|G|G|T|C|C|T|A|T|G|A|C|T|G|T|G|G|G|C|C|A|C|T|C|T|C|T|G|A|G|A|C|C|G|C|A|A|T|G|C|G 994
QY 261 L|E|U|P|H|E|I|S|T|H|R|P|R|O|T|H|R|G|L|Y|H|I|S|I|L|E|G|L|U|A|L|E|U|R|G|L|N|G|L|U|I|L|E|A|R|N|A|L|A 280
DB 995 C|T|T|T|T|C|A|G|C|T|C|T|C|T|C|A|G|A|G|C|A|T|G|A|A|G|C|C|T|C|G|G|C|A|G|G|A|T|G|C|A|T|G|C|G|C|C 1054
QY 281 A|R|G|P|R|G|L|N|C|Y|R|P|R|O|V|A|L|G|L|Y|L|E|U|A|R|N|T|H|L|E|U|A|R|P|R|O|S|E|R|I|L|E|A|R|G|L|Y|R|G|L|U 300
DB 1055 C|A|C|C|C|C|A|G|G|C|C|C|G|G|G|C|C|T|T|A|C|A|C|C|T|G|C|T|T|C|C|C|A|G|A|T|C|A|C|C|G|A|G|A|A 1114
QY 301 V|A|L|A|G|L|U|G|L|U|Y|R|S|E|R|P|R|O|T|P|R|A|L|Y|R|L|E|U|S|E|R|C|Y|S|G|L|Y|H|I|S|A|L|I|S|G|L|Y|R|H|S 320
DB 1115 G|T|G|T|G|A|A|G|A|G|A|G|C|C|T|G|G|G|A|T|A|C|T|G|A|G|T|G|G|G|C|A|T|G|G|C|A|G|G|G|T|A|C|C 1174
QY 321 A|R|N|T|P|R|G|L|Y|H|I|S|A|R|G|S|E|R|A|R|P|R|H|R|G|L|U|A|A|R|G|L|U|R|G|L|U|C|Y|R|P|R|O|T|C|Y|R|A|R|G|T|H|R 340
DB 1175 C|A|C|T|G|G|G|G|C|A|T|G|G|G|C|A|C|A|G|G|A|G|C|C|A|G|A|G|G|G|A|G|G|T|C|C|C|A|G|T|C|A|C|A|G|T|C|A|C|A|C|T 1234
QY 341 V|A|L|G|L|Y|R|P|R|O|T|Y|R|A|L|P|R|O|L|E|U|T|P|R|L|E|U|G|L|Y|C|Y|S|G|L|U|A|G|L|Y|P|H|E|R|Y|R|A|L|A|R|P|A|L|G|L|Y 360
DB 1235 G|T|G|G|G|C|C|C|A|G|C|C|T|C|T|G|G|C|T|G|G|G|C|T|G|A|G|C|A|G|A|T|T|T|A|T|G|T|G|A|N|N|N|N|N 1294
QY 361 P|R|O|P|R|O|T|H|R|I|S|A|L|P|H|E|R|P|R|O|C|Y|S|G|L|Y|H|I|S|A|L|Y|C|Y|S|E|R|G|L|U|Y|R|S|E|R|A|L|Y|R|Y 380
DB 1295 N|N|N|N|N|N|C|A|T|G|T|T|T|C|A|C|C|C|T|G|G|G|C|A|G|T|G|T|T|C|A|G|A|G|A|G|T|T|G|C|A|A|G|T|A|C 1354
QY 381 T|P|R|S|E|R|G|L|N|L|E|P|R|O|L|E|U|R|P|R|O|H|I|S|G|L|Y|T|H|H|I|S|A|L|P|H|E|I|S|A|L|A|C|Y|R|P|R|O|P|R|E|C|Y|S 400

```

```

DB 1355 T|G|T|T|C|C|A|A|T|C|C|A|T|G|C|C|C|A|C|G|A|G|C|A|G|C|G|T|T|C|A|G|C|G|C|G|T|C|G|T|C|T|G 1414
QY 401 A|A|T|H|R|G|L|U|V|A|L|G|L|U|G|L|N|A|R|G|L|Y|L|E|U|S|E|R|I|L|E|P|H|E|R|G|L|N|G|L|Y|R|P|R|O|L|E|A|R|P 420
DB 1415 G|C|C|A|C|C|A|C|T|G|T|T|G|T|G|A|C|A|G|A|C|T|C|A|T|C|A|A|T|G|A|T|T|T|C|A|A|G|T|C|A|G|T|G|A|C 1474

RESULT 13
AF302504
LOCUS AF302504 1734 bp mRNA linear ROD 05-JUL-2001
DEFINITION Mus musculus pellino 2 (Pell2) mRNA, complete cds.
ACCESSION AF302504
VERSION AF302504.1 GI:10242356
KEYWORDS Mus musculus (house mouse)
SOURCE Mus musculus
ORGANISM Mus musculus
REFERENCE 1 (bases 1 to 1734)
AUTHORS Resch, K., Jockusch, H. and Schmitt-John, T.
TITLES Assignment of homologous genes, Pell1/Pell1 and Pell2/Pell2, for the Pelle adaptor protein Pellino to mouse chromosomes 11 and 14 and human chromosomes 2p13.3 and 14q21, respectively, by physical and radiation hybrid mapping
JOURNAL Cytogenet. Cell Genet. 92 (1-2), 172-174 (2001)
PUBMED 11306823
REFERENCE 2 (bases 1 to 1734)
AUTHORS Resch, K., Jockusch, H. and Schmitt-John, T.
TITLES Direct Submision
JOURNAL Submitted (05-SEP-2000) Developmental Biology and Molecular Pathology, University of Bielefeld, Universitaetsstrasse 25, Bielefeld, NRW 33615, Germany
FEATURES
source
1..1734
/organism="Mus musculus"
/mol_type="mRNA"
/strain="C57BL/6J"
/db_xref="taxon:10090"
1..1734
/gene="Pell2"
218..1477
/gene="Pell2"
/note="similar to Drosophila melanogaster pellino"
/codon_start=1
/product="pellino 2"
/protein_id="AA015392.1"
/db_xref="GI:10242357"
/translation="MPSPEQSPAPNKRPKYRRLVULGYNALPNDGRKKRSFT
LKRIVASGVKPSSTIMHVSPTQASRAISRKHHSISYLSRQTVVETHTDQDMF
QVGRSTESPIDIVVDIVSFGQNEAQTOSTSRACRIVDRNRPYARIPAAFD
SSKNIFLGEKAKMKNPQDGHMDGLTNGLVNHPQGFPEESQPGVWRISVCGDVT
LRETSAAOGRGLVSEETNVLQDGLIDLCATVLMRTDGLFHAFTOKHIEAROEI
NBARPCPYGLNTLAPESTINREKVEERQPMAYLSCGHVGYHMHGRSDTEANREC
PMQTVGPVPLMLGCEAFYIYXXXXXHVLPFCGVSEKSAKTYSLPLPHGTAP
HAACPFCATQVLVBQNCIKLIFQGPVD"

ORIGIN
Alignment Scores:
Pred. No.: 8 646-153 Length: 1734
Score: 2111.50 Matches: 388
Percent Similarity: 95.2% Conservative: 12
Best Local Similarity: 92.4% Mismatches: 19
Query Match: 92.2% Indels: 1
DB: 9 Gaps: 1
US-10-041-030-4 (1-420) x AF302504 (1-1734)
QY 1 MetPheSerProGlnGlnGlnGlnHisCybAlaProAsnLysGlnUroVAlaYsYrGly 20
DB 218 ATGTTTCCCGGGGCGAGAGAAACCCAGCCGCCCAACAGAGCGGTGAATACAGG 277
QY 21 G|U|L|E|U|V|A|L|E|U|G|L|Y|T|Y|R|A|N|G|L|Y|A|L|E|U|P|R|O|A|N|G|L|Y|A|S|P|R|G|L|Y|R|G|L|Y|S 40

```

```

|||||
278 GACGTGTCGTCGGGGTCAATGTCCTTACCTATGTCAGACGGGCAAGAA 337
Qy |||||
41 SerArgPheAlaLeuTyrLysArgProLysAlaAsnGlyValLysProSerThrValHis 60
Db |||||
338 AGCAGATTATCCCTTATTAAGCGGACCTACGCGGAGTGTGTCAAAACCCACACATC 397
Qy |||||
61 ValIleSerThrProGlnAlaSerLysAlaIleSerCysLysGlyGlnHisSerIleSer 80
Db |||||
398 ATGTCCTCCACACACAGCGGTCACAGCCATCAGCTCCAGAGACATCAGCATTCG 457
Qy |||||
81 TyrThrLeuSerArgAsnGlnThrValValGlnTyrThrHisAspLysAspThrAsp 100
Db |||||
458 TACACGTTGTTCACGAGCCAGACGCGTAGTGTGAGTACACACGATTAAGACACAGAC 517
Qy |||||
101 MetPheGlnValGlyArgSerThrGlnSerProLysPhePheValValThrAspThrIle 120
Db |||||
518 ATGTTTCAGGTGGGCGAGTCAAGAAAGCCCATTCCTGTCGTCACAGACGCGTT 577
Qy |||||
121 SerGlySerGlnAsnThrAspGlnValGlnIleThrGlnSerThrIleSerArgPheAla 140
Db |||||
578 TCCGGGGCGTCAGAAAGAAAT---GCCCATATCACACAGACCATCTAGGTTCCGA 634
Qy |||||
141 CysArgIleValCysAspArgAsnGlnProTyrThrAlaArgIlePheAlaIleGlyPhe 160
Db |||||
635 TGCAGATCGTGTGTGACAGAAACGACCATACACAGCATATTCGCGCGCAGATTC 694
Qy |||||
161 ArgSerSerLysAsnIlePheLeuGlyGlnLysAlaAlaIleTyrPheAspProAspGly 180
Db |||||
695 GATTCCTTCCAAAATATCTTCTTGGAGAAAGACAAATGGAATGGAATGGAATGGA 754
Qy |||||
181 HisMetAspGlyLeuThrThrAsnGlyValIleValValMetHisProArgGlyGlyPheThr 200
Db |||||
755 CACATGATGATGATCACTACCAACGCTCTCTGTGATGACCCGCAAGAGGCTTACC 814
Qy |||||
201 GlnGlnSerGlnProGlyValITPArgGlnIleSerValCysGlyAspValTyrThrLeu 220
Db |||||
815 GAGGAATCCAGCGCTGAGTCTGAGAGAGATCTCTGTCGTGGGATGATACACCTTG 874
Qy |||||
221 ArgGlnThrArgSerAlaGlnGlnValGlyLysValGlnSerGlnThrAsnValLeu 240
Db |||||
875 CGAGAGCCAGGTCGCGCCAGAGAGGGAAGCTGTGGAAGTGAACCAACGTCCTG 934
Qy |||||
241 GlnAspGlySerLeuIleAspLeuCysGlyValIleThrLeuTyrPArgThrAlaAspGly 260
Db |||||
935 CAGAGCGGCTCCCTCATTTGACTGTGTGGGCCCATCTCTCTGGAACACCGCATGCG 994
Qy |||||
261 LeuPheHisThrProThrGlnLysHisIleGlnAlaLeuArgGlnIleAsnAlaAla 280
Db |||||
995 CTTTTCACGCTCTACTACAGAGCATAGAGCCCTCGGCGAGATCAATGACAGCC 1054
Qy |||||
281 ArgProGlnCysProValGlyLeuAsnThrLeuAlaPheProSerIleAsnArgLysGln 300
Db |||||
1055 CGAGCCAGAGGCCCCGGGGCCTTAAACCTTGCTTCCCGACATCAACCGGAAGAA 1114
Qy |||||
301 ValValGlnGlnLysGlnProTyrPalaTyrLeuSerCysGlyHisValHisGlyTyrHis 320
Db |||||
1115 GTGTGTGAAGAAGAACGCGCTGTGGCATCTGAGCTGCGGCATGTGACGCGCTACAC 1174
Qy |||||
321 AsnITPArgLysHisArgSerAspThrGlnAlaAsnGlnLysGlnCysProMetCysArgThr 340
Db |||||
1175 CACTGGGGCGCATCGAGACGACACGGAAGCCACGAGGGAGTGTCCCATGTGCATGACT 1234
Qy |||||
341 ValGlyProTyrValProLeuTyrPheGlyCysGlnAlaGlyPheTyrValAspAlaGly 360
Db |||||
1235 GTGGGCCCCCAGCGCTCTCTGTGCTGTGAGTGTGACGAGATTTTATGTCGANNNNNN 1294
Qy |||||
361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGlnLysSerAlaLysTyr 380
Db |||||
1295 NNNNNNNCTATGTTTTCACCCCTGCGGGGACGTCTGTTTCAGGAAGTGTGCCAAGTAC 1354
Qy |||||
381 TrpSerGlnIleProLeuProHisGlyThrHisAlaPheHisAlaAlaCysProPheCys 400

```

```

Db 1355 TGTCCAGATCCCATCGCCCGCCAGGACGACGCGTTTTCACGCGCTGTCCCTTCG 1414
Qy 401 AlaThrGlnLeuValGlyGlnGlnAsnCysIleLysLeuIlePheGlnGlyProIleAsp 420
Db 1415 GCCACGAGCTGTGTGTGAAAGAACTGCATCAATGATTTTCCAAAGTCCAGTGGAC 1474

RESULT 14
LOCUS BC075973
DEFINITION Danio rerio p011no homolog 2 (Drosophila), mRNA (CDNA clone
ACCESSION BC075973
VERSION BC075973.1 GI:49903830
KEYWORDS MGC.
SOURCE Danio rerio (zebrafish)
ORGANISM Danio rerio
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
Cypriniformes; Cyprinidae; Danio.
REFERENCE
AUTHORS Strausberg, R.L., Feingold, E.A., Grouse, L.H., Derge, J.G.,
Klausner, R.D., Collins, F.S., Wagner, L., Shenmen, C.M., Schuler, G.D.,
Altschul, S.F., Zeeberg, B., Buetow, K.H., Schaefer, C.F., Bhat, N.K.,
Hopkins, R.F., Jordan, H., Moore, T., Max, S.I., Wang, J., Hsieh, F.,
Diatchenko, L., Marusina, K., Farmer, A.A., Rubin, G.M., Hong, L.,
Stapleton, M., Soares, M.B., Bonaldo, M.F., Casavant, T.L.,
Scheltz, T.E., Brownstein, M.J., Usdin, T.B., Tashy, T.S.,
Carninci, P., Prange, C., Raha, S.S., Loquellano, N.A., Peters, G.J.,
Adams, R.D., Mullan, S.J., Bosak, S.A., McEwan, P.J.,
McKernan, K.J., Malek, J.A., Gunaratne, P.H., Richards, S.,
Worley, K.C., Hale, S., Garcia, A.M., Gay, L.J., Hui, S.W.,
Villalon, D.K., Muzny, D.M., Sodergren, E.J., Lu, X., Gibbs, R.A.,
Fahy, J., Helton, E., Kettman, M., Madan, A., Rodriguez, S.,
Sanchez, J., Whitting, M., Madan, A., Young, A.C., Shcherbakov, Y.,
Boutard, G., Blakeley, R.W., Touchman, J.W., Green, E.D.,
Dickson, M.C., Rodriguez, A.C., Grimwood, J., Schmutz, J., Myers, R.M.,
Butterfield, Y.S., Krzywinski, M.I., Skalska, U., Small, D.B.,
Schnerch, A., Schein, J.E., Jones, S.J. and Marra, M.A.
Mammalian Gene Collection Program Team
Generation and initial analysis of more than 15,000 full-length
human and mouse cDNA sequences
Proc. Natl. Acad. Sci. U.S.A. 99 (26), 16899-16903 (2002)
12477932
2 (bases 1 to 2474)
JOURNAL NIH MGC Project
PUBMED Direct Submission
AUTHORS Submitted (06-JUL-2004) National Institutes of Health, Mammalian
JOURNAL Gene Collection (MGC), Bethesda, MD 20892-2590, USA
REMARK NIH-MGC Project URL: http://mgc.ncl.nih.gov
COMMENT Contact: MGC help desk.
Email: gcgabs-remail.nih.gov
Tissue Procurement: Len Zon, Harvard
CDNA Library Preparation: Open Biosystems
DNA Sequencing by: The I.M.A.G.E. Consortium (LNL)
Center, Stanford University School of Medicine, Stanford, CA 94305
Web site: http://www-shgc.stanford.edu
Contact: (Dickson, Mark) mcd@paxil.stanford.edu
Dickson, M., Schmutz, J., Grimwood, J., Rodriguez, A., and Myers,
R. M.
Clone distribution: MGC clone distribution information can be found
through the I.M.A.G.E. Consortium/LNL at: http://image.llnl.gov
Series: IRK Plate: 179 Row: m Column: 7
This clone was selected for full length sequencing because it
passed the following selection criteria: matched mRNA gi: 49903830.
location/Qualifiers
1..2474
/organism="Danio rerio"
/mol_type="mRNA"
/db_xref="taxon:7955"
/clone="MGC:92268 IMAGE:7053197"
FEATURES
Source

```



```

/tissue type="Whole, adult zebrafish"
/clone_id="N1H_ZGC_7"
/lab host="DH10B"
/notes="Vector: pExpress1"
1..2474
/gene="pel12"
/notes="synonym: zgc:92268"
/db_xref="GeneID:436889"
/db_xref="ZFIN:ZDB-GENE-040718-360"
218..1504
/gene="pel12"
/codon_start=1
/product="pellino homolog 2"
/protein_id="AAH75973.1"
/db_xref="GI:49903831"
/db_xref="GeneID:436889"
/db_xref="ZFIN:ZDB-GENE-040718-360"
/translacion="MESPQGEHCAPKEPVKGLVVLGNSLPNGDRGRKSRPA
LYRKAKGVKPSVHLANTPOASKAVNGSHISYLSRNOTVVEYSHDQDMF
OIGRSTGPIDVVTIVSGGSEERTPIOSTISRACVCEKRNPPYARIPAFGE
DSKNIPLGKRAKMKRPGDHMGDTTNGVLNMRPGFTBESKGVREISVCGDY
TLBETNSAOTRKLVSESNVLDGSLVDLCAATLMLRTADGLFHTPTQKHLARQE
LNARPCQVGLNTLAPSPMQRCSALSNVQEDKQPMWYIACGHVAGHDHGRSE
RESNAORECEMCRVGPVPLVLMGCEBAPFYVDGAPFHVFPVCGHVCSEKSTYMAEI
PLPHGTAFPAACPFCAQTQNLTKMAKLI.FGCPID"

```

ORIGIN

Alignment Scores:

```

Pred. No.: 1.67e-144 Length: 2474
Score: 2006.00 Matches: 361
Percent Similarity: 90.9% Conservative: 28
Best Local Similarity: 84.3% Mismatches: 31
Query Match: 87.6% Indels: 8
DB: 5 Gaps: 2

```

US-10-041-030-4 (1-420) x BC075973 (1-2474)

```

QY 1 MetPheSerProGluGluGluGluHisCysAlaProbaNlySGluProValIlyrTgLy 20
Db 218 ATGTTTTCACCGGGTCAAGAGAGACCTGTGCCCAAAAGCCAGTAAGTACGGG 277
QY 21 GluLeuValIalleuGlyTyraNGlyAlaLeuProbaNlyAaPaRgGlyArGaRgLy 40
Db 278 GAATGATGTTTGAATATGATATGCTCACTCCCAATGCTGACAGAGGCGAAGAAA 337
QY 41 SerArgPheAlaLeuTyTyraNGlyProLySaAlaNGlyValIlyrProSerThrValHis 60
Db 338 AGCGGTTTCCCTTCAAGAGAGCCAAAGCCAAATGGGGTCAAGCCAGACTGACAC 397
QY 61 ValIleSerThrProGluAlaSerLySaAlaIleSerCysLySGlyGluHisSerIleSer 80
Db 398 ATCTTAAACACACCAAGCCAGCCAGGCAAGTAACTGAAGGCGCAGACAGATCTCC 457
QY 81 TyrThrLeuSerArgaNGluInThrValIValIgluTyrThrHisAaPlyAaPThrAsp 100
Db 458 TACACGCTGTCCAGAAACCCAGACCGTGTGTGAGTACAGCCATGACAAAGACACGAC 517
QY 101 MetPheGluValGlyArGaSerThrGluSerProIleAsPheValIValThraSPThrIle 120
Db 518 ATGTTCCAGATCGGCGCATCGAGGAGGCGCCGATTCATCTTGTGTGAGTCACTGAG 577
QY 121 SerGlySerGluInThraSPheGluAlaGluIleThrGluInSerThrIleSerArgPheAla 140
Db 578 TCAAGGGGAGAGAGAGTGAAGACACCCATCACTCAAAAGCACCATCTCCCGCTTGC 637
QY 141 CysArgIleValIValSaPaRgAaNGluProTyThrAlaRglIlePheAlaAlaGlyPhe 160
Db 638 TGCAGGAGTGGTGTGTGAGTCAACCCATCAACCCAGAACTTTCGCTGACGAGCTTC 697
QY 161 AspSerSerLySaNGlylePheLeuGlyGluLySaAlaIleAlaIlyrTplyAaNGly 180
Db 698 GATTCCTTCAAGATATATTTCTCTCGAGAAAAGCAAGCAATGGAAGATCCAGATGCT 757

```

```

QY 181 HisMetAspGlyLeuThrThrAnGlyValIleuValMetHisSPRoARgGlyGlyPheThr 200
Db 758 CACATGACCGTTTGACTACCAATGGGCTCTATGATGACACCGCGGGCGGTTTACA 817
QY 201 GluGluSerGlnProGlyValIValPaRgGluIleSerValCysGlyAaPValIlyrThrLeu 220
Db 818 GAGGAATCCAAAGCTGTGTGTGAGGAGATATCAAGTGTGTGAGATGTTTACACACTG 877
QY 221 ArgGluThrArgSerAlaGluNGluValGlyLyLeuValIleGluSerGluInThraNGluLeu 240
Db 878 AGAGAAACACGCTCTGTCTCAAACTCGACGAAATGTGTGAGAGTGAAGTAAAGCTTCTG 937
QY 241 GluAspGlySerLeuIleAaPLeuCysGlyValIValThrLeuLeuThrPaRgThrAlaAspGly 260
Db 938 CAGACGCGCTCTGTGTGATCCTTTGTGAGCAACGCTGTGTGCGCACCGCTGACGCG 997
QY 261 LeuPheHisThrProThrGluNGluHisIleGluAlaLeuArgGluIleAsnAlaAla 280
Db 998 CTTTCCACACCCCCACACAGAAAGCACTGAGAGGCTTGTGCGCAGAGCTGAAGCGAC 1057
QY 281 ArgProGluCysProValIleGluLeuInThrLeuAlaPheProSerIleAsnArg----- 298
Db 1058 CGGCCCCAGTGCCTGTGTGCGCTCAACACACTGGCTTCCACAGATGCAACGCTGCAGT 1117
QY 299 -----LysGluValValIleGluValGluGluProThrPaIaTyrlLeuSerCys 313
Db 1118 GTGCCCTGACGACACGCTCCTGCCAGAGAAAGACAAAGCCGTGGGTCTACCTTGCCTGT 1177
QY 314 GlyHisValHisGlyTyThrHisAaNGlyProGlyHisArgSerAaPThrGluAlaAsn---Glu 332
Db 1178 GTTCATGTTTATATGCTGACACACACTGGGCGCACCATGACAGGCGGATCCATGCCAA 1237
QY 333 ArgGluCysProMetCysArgThrValIleGlyProTyValIProleuThrLeuGlyCysGlu 352
Db 1238 AGAGAGTGCCTCAATGTGACAGGCGCGGTGGTCCCTAATGTGCCACTGTGGCTCGGATGAA 1297
QY 353 AlaGlyPheTyThrValaPheAlaGlyProProThrHisAlaPheThrProCysGlyHisVal 372
Db 1298 CAGACCTTCTACGTACACACTGTGTGCGCAACACAGTCTTCTGTACTGTGTGCGCAATGTG 1357
QY 373 CysSerGluCysSerAlaLeuTyTyraNGlyProLeuProIleSerGluThrHisAla 392
Db 1358 TGCCTGAAAGTCCACCGCTATTTGGGACAGAAATCTTTGCCAATGACACCAAGCT 1417
QY 393 PheHisAlaAlaCysProPheCysAlaThrGluLeuValIleGluGluInSerCysIleLy 412
Db 1418 TTCCAGGCTGCTGTCTTCTGTGCTACCTCACTCAACCTGACACAAAGTGGGCCAAG 1477
QY 413 LeuIlePheGluNGlyProIleAsp 420
Db 1478 CTCATCTTTCAGGGGCTATAGAC 1501

```

RESULT 15

AR483566 1257 bp DNA linear PAT 14-MAY-2004

LOCUS AR483566 Sequence 3 from patent US 6703487.

DEFINITION AR483566

ACCESSION AR483566

VERSION AR483566.1 GI:47246650

KEYWORDS

SOURCE

ORGANISM

REFERENCE

AUTHORS

TITLE

JOURNAL

FEATURES

source

ORIGIN

Unclassified.
1 (bases 1 to 1257)
Bird, T.A. and Cosman, D.J.
Human pellino polypeptides
patent: US 6703487-A 3 09-MAR-2004;
ImmuneX Corporation; Seattle, WA;
WOX;

Location/Qualifiers
1..1257
/organism="unknown"
/mol_type="genomic DNA"

Alignment Scores:

| | | | |
|------------------------|-----------|---------------|------|
| Pred. No.: | 5,34e-138 | Length: | 1257 |
| Score: | 1917.00 | Matches: | 342 |
| Percent Similarity: | 90.5% | Conservative: | 38 |
| Best Local Similarity: | 81.4% | Mismatch: | 38 |
| Query Match: | 83.7% | Indels: | 2 |
| | | Gaps: | 1 |

US-10-041-030-4 (1-420) x AR483566 (1-1257)

```

QY      1 MetPheSerProGlnGlnGluHisCysAlaProAsnIysGluProValIlyrTyGly 20
DB      1 ATGTTTCTCTGATCAAGAAATCAT-----CCATCTAAGACACAGTAAATATGTT 54
QY      21 GluLeuValIleuGlyTyraGlnIyAlaLeuProAsnIysAspArgIyAaArgIy 40
DB      55 GAACTCATTTGCTTAAAGATTAATGATCTCTCCAAACGATGTAAGAGAGAGGAAA 114
QY      41 SerArgPheAlaLeuTyIysArgProIyAlaAsnIyValIyPProSerThrValHis 60
DB      115 AGTAGTTTCTTGTTTAAAGACCTAAGGCATTAAGGATGGAAGCCACACATGTTCAT 174
QY      61 ValIleSerThrProGlnAlaSerIyAlaIleSerCysIysGlyGlnHisSerIleSer 80
DB      175 ATTGCTTGTACTCTCAGGCTGCAAGAGCAATAAAGCAAAAGACCATAGCATATCA 234
QY      81 TyrThrIleuSerArgAsnGlnThrValIValIyGluTyThrHisAspIyAspThrAsp 100
DB      235 TATACCTTATCTCGGCGCCCAAGACTGTGTGTGAATATCTCATGACACCAACCGAT 294
QY      101 MetPheGlnValIyArgSerThrGluSerProIleAspPheValIThrAspThrIle 120
DB      295 ATGTTTCAGATTGGCCGCTGACCTGAAGCCCATTTGTTGTAATCTGAACACCGTT 354
QY      121 SerGlySerGlnAsnThrAspGluAlaGlnIleThrGlnSerThrIleSerArgPheAla 140
DB      355 CCTGGAAGTCAAGTAAATTCTGATACACAGTCAAGTCAAAAGCACTATATCAAGATTGCC 414
QY      141 CysArgIleValCysAspArgAsnGluProTyThrIleAlaGlyIlePheAlaIyPhe 160
DB      415 TGGAGAACTATATGTCAGGAATCTCCCTTTACACACGGAATTATGCTGCGAGGTTT 474
QY      161 AspSerSerIyAsnIlePheLeuGlyIyAlaIyAspIyAspProAspGly 180
DB      475 GACTCATCAAAAACATCTTTCTTGGGGAAGAGCTCCAAATGGAAGACATCAGATGA 534
QY      181 HisMetAspGlyLeuThrIleAsnGlyValIleuValIleHisProArgIyGlyPheThr 200
DB      535 CAGATGATGCTTGAACCACTAATGATGTTCTTGTGATGATCCACGCAATGGGTTACA 594
QY      201 GluGluSerGlnProGlyValITrArgGluIleSerValCysGlyAspValITyThrIleu 220
DB      595 GAAAGACTCCAAAGCTGGAATATGGAAGAAATATCGGTGTGGAAATGTAATTAAGCTTA 654
QY      221 ArgGluThrArgSerAlaGlnIleArgIyIyLeuValIyGluSerGluThrAsnValIleu 240
DB      655 CGTGAACCAAGATGCGTCTCAGCAGAGAGAAAATGCTGGAATTAACCAATCACTTA 714
QY      241 GluAspGlySerIleuIleAspLeuCysGlyAlaThrIleuLeuITrArgThrAlaAspGly 260
DB      715 CAAGATGCTCGTAAATTAATGACCTGTGTGCAACATTTATGCGCTACGCAAGGC 774
QY      261 LeuPheIleThrProThrGlnIyIyHisIleGluAlaLeuArgGlnIyIleAsnAlaIle 280
DB      775 CTTTCCCACTCTTACCTGGAAGCATTTAAGACGGAATCAATGCAAGCA 834
QY      281 ArgProGlnCysProValIyIyLeuAsnThrIleuAlaPheProSerIleAsnArgIyGlu 300
DB      835 CGAAGCTAGTCCCTGTAGGCTTCAACACATGACATTTCTTATGTAAGAGGAAGAC 894
QY      301 ValIyGluGlyIyGlnProITrPalaITyIleuSerCysGlyHisValIyGlyIyHis 320
DB      |||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::|||||:::

```

```

DB      895 GTTGTAGATGAAAAACAACATGGGTATATCTAACTGGGCGCATGTACATGCTATCAT 954
QY      321 AsnTrpGlyHisArgSerAspThrGluAlaAsnGluAArgGluCysProMetCysArgThr 340
DB      955 AACTGGGGAACAAAGAAAGAACTGATGGCAAGATCGTAATGTCTTATGTATAGTCT 1014
QY      341 ValIyProTyValProLeuITrLeuGlyCysGluAlaGlyPheTyValAlaPalaGly 360
DB      1015 GTTGTCCCTATGTTCTGTGTGCTTGGATGTGAAGCTGGAATTTATGTGACCGCGCG 1074
QY      361 ProProThrHisAlaPheThrProCysGlyHisValCysSerGluIySerAlaIyTy 380
DB      1075 CCTCCAAACCATGCGTTTAAAGCCGTGTGGCATGTGTGTCAAAAAAGCAACTGCTAT 1134
QY      381 TrpSerGlnIleProLeuProHisGlyITrHisAlaPheHisAlaIyCysProPheCys 400
DB      1135 TGGTCCCAATCCCACTTCTCTCAAGGATCTCATCTTTCATGACACGCTGTCTTTGT 1194
QY      401 AlaThrGlnLeuValIyGluGlnAsnCyHisIyLeuIlePheGlnGlyProIleAsp 420
DB      1195 GCACATCAGTTGGCTGTGAACAAGGCTATCATCAGACTTATTTTCAAGGACCTTAAGAC 1254

```

Search completed: March 2, 2006, 07:33:41
 Job time : 6959 secs

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2006, 21:01:56 ; Search time 20 Seconds
(without alignments)
418.898 Million cell updates/sec

Title: US-10-041-030-4

Perfect score: 2290
Sequence: 1 MFSFGQEHKAPNKPKYKVG.....ATQVGEQNCIKIFQGPID 420

Scoring table: BLOSUM62
Gapop 10.0, Gapext 0.5

Searched: 133702 seqs, 19947517 residues

Total number of hits satisfying chosen parameters: 133702

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database: Published Applications AA New:

1: /cgn2_6/ptodaca/1/pubpaa/US06_NEW_PUB.pep:*
2: /cgn2_6/ptodaca/1/pubpaa/US07_NEW_PUB.pep:*
3: /cgn2_6/ptodaca/1/pubpaa/US07_NEW_PUB.pep:*
4: /cgn2_6/ptodaca/1/pubpaa/PCR_NEW_PUB.pep:*
5: /cgn2_6/ptodaca/1/pubpaa/US09_NEW_PUB.pep:*
6: /cgn2_6/ptodaca/1/pubpaa/US10_NEW_PUB.pep:*
7: /cgn2_6/ptodaca/1/pubpaa/US11_NEW_PUB.pep:*
8: /cgn2_6/ptodaca/1/pubpaa/US60_NEW_PUB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----------------------|--------------------|
| 1 | 1482.5 | 64.7 | 406 | 7 US-11-072-512-3928 | Sequence 3928, Ap |
| 2 | 90 | 3.9 | 800 | 7 US-11-024-959-511 | Sequence 511, App |
| 3 | 90 | 3.9 | 984 | 6 US-10-995-561-629 | Sequence 629, App |
| 4 | 89 | 3.9 | 452 | 7 US-11-087-099-7273 | Sequence 7273, App |
| 5 | 88 | 3.9 | 800 | 7 US-11-052-554-282 | Sequence 282, App |
| 6 | 88 | 3.8 | 1345 | 7 US-11-024-959-471 | Sequence 471, App |
| 7 | 87.5 | 3.8 | 1121 | 7 US-11-052-554-282 | Sequence 459, App |
| 8 | 87 | 3.8 | 320 | 7 US-11-087-099-7278 | Sequence 7278, App |
| 9 | 87 | 3.8 | 2105 | 7 US-11-052-554-173 | Sequence 173, App |
| 10 | 86.5 | 3.8 | 266 | 7 US-11-072-512-3891 | Sequence 3891, App |
| 11 | 86 | 3.8 | 759 | 7 US-11-149-003-22 | Sequence 22, App |
| 12 | 86 | 3.8 | 1057 | 7 US-11-149-003-6 | Sequence 6, App |
| 13 | 86 | 3.8 | 1192 | 7 US-11-149-003-18 | Sequence 18, App |
| 14 | 86 | 3.8 | 1207 | 7 US-11-149-003-20 | Sequence 20, App |
| 15 | 86 | 3.8 | 1251 | 7 US-11-149-003-16 | Sequence 16, App |
| 16 | 86 | 3.8 | 1342 | 7 US-11-149-003-24 | Sequence 24, App |
| 17 | 86 | 3.8 | 1477 | 7 US-11-149-003-8 | Sequence 8, App |
| 18 | 86 | 3.8 | 1512 | 7 US-11-149-003-10 | Sequence 10, App |
| 19 | 86 | 3.8 | 1535 | 7 US-11-149-003-14 | Sequence 14, App |
| 20 | 86 | 3.8 | 1570 | 7 US-11-149-003-12 | Sequence 12, App |
| 21 | 86 | 3.8 | 1593 | 7 US-11-149-003-4 | Sequence 4, App |
| 22 | 86 | 3.8 | 1628 | 7 US-11-149-003-2 | Sequence 2, App |
| 23 | 86 | 3.8 | 2176 | 7 US-11-193-561-25 | Sequence 25, App |
| 24 | 86 | 3.8 | 2176 | 7 US-11-193-771-25 | Sequence 25, App |
| 25 | 86 | 3.8 | 2176 | 7 US-11-193-789-25 | Sequence 25, App |

| | | | | | |
|----|----|-----|------|---------------------|-------------------|
| 26 | 86 | 3.8 | 2176 | 7 US-11-193-806-25 | Sequence 25, App |
| 27 | 86 | 3.8 | 2176 | 7 US-11-193-857-25 | Sequence 25, App |
| 28 | 86 | 3.8 | 2217 | 7 US-11-193-561-38 | Sequence 38, App |
| 29 | 86 | 3.8 | 2217 | 7 US-11-193-771-38 | Sequence 38, App |
| 30 | 86 | 3.8 | 2217 | 7 US-11-193-789-38 | Sequence 38, App |
| 31 | 86 | 3.8 | 2217 | 7 US-11-193-806-38 | Sequence 38, App |
| 32 | 86 | 3.8 | 2217 | 7 US-11-193-857-38 | Sequence 38, App |
| 33 | 86 | 3.8 | 2223 | 7 US-11-193-561-2 | Sequence 2, App |
| 34 | 86 | 3.8 | 2223 | 7 US-11-193-771-2 | Sequence 2, App |
| 35 | 86 | 3.8 | 2223 | 7 US-11-193-789-2 | Sequence 2, App |
| 36 | 86 | 3.8 | 2223 | 7 US-11-193-806-2 | Sequence 2, App |
| 37 | 86 | 3.8 | 2223 | 7 US-11-193-857-2 | Sequence 2, App |
| 38 | 86 | 3.8 | 2296 | 6 US-10-995-561-633 | Sequence 633, App |
| 39 | 86 | 3.8 | 2296 | 7 US-11-193-561-23 | Sequence 23, App |
| 40 | 86 | 3.8 | 2296 | 7 US-11-193-771-23 | Sequence 23, App |
| 41 | 86 | 3.8 | 2296 | 7 US-11-193-789-23 | Sequence 23, App |
| 42 | 86 | 3.8 | 2296 | 7 US-11-193-806-23 | Sequence 23, App |
| 43 | 86 | 3.8 | 2296 | 7 US-11-193-857-23 | Sequence 23, App |
| 44 | 86 | 3.8 | 2330 | 7 US-11-193-561-21 | Sequence 21, App |
| 45 | 86 | 3.8 | 2330 | 7 US-11-193-771-21 | Sequence 21, App |

ALIGNMENTS

RESULT 1
US-11-072-512-3928
Sequence 3928, Application US/11072512
Publication No. US2006002945A1
GENERAL INFORMATION:
APPLICANT: ISOGAI, TAKAO
APPLICANT: SUGIYAMA, TOMOYASU
APPLICANT: OTSUKI, TETSUJI
APPLICANT: WAKAMATSU, AI
APPLICANT: SATO, HIROYUKI
APPLICANT: ISHII, SHIZUKO
APPLICANT: YAMAMOTO, JUN-ICHI
APPLICANT: ISONO, YUUKO
APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHITO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOKU
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: Novel full length cDNA
FILE REFERENCE: 084315-0191
CURRENT APPLICATION NUMBER: US/11/072,512
CURRENT FILING DATE: 2005-03-07
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: US 60/350,978
PRIOR APPLICATION NUMBER: JP 2001-379298
PRIOR FILING DATE: 2001-11-05
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3928
LENGTH: 406
TYPE: PRT
ORGANISM: Homo sapiens
US-11-072-512-3928
Query Match 64.7%; Score 1482.5; DB 7; Length 406;
Best Local Similarity 65.6%; Pred. No. 4.8e-130;
Matches 275; Conservative 36; Mismatches 65; Indels 43; Gaps 4;
Cy 3 SFGQEHKAPNKPKYKVGELVLYGYNALPNDGRGRKSRFALYKPKANGVSPVTHVI 62
Db 30 SFGQEHKAPNKPKYKVGELVLYGYNALPNDGRGRKSRFALYKPKANGVSPVTHVI 62
Cy 63 STFOAKSKALSKQKHSHISTLSRNQTVVVEYTHDQTDMEFQVGRSTESPIDFVYDTISG 122

```

Db      49  STEPLVSKALSNRGQHSISVITLRSRSHVEIYETHDSTDMDQJQRSTENMIDPVTDTSPG 108
Oy      123  SONTDEAOITGSTISIFRACIVCDNEPPYTARIIPAAGPSSKNIPIGEKAKKKNPDGHH 1822
Db      1109  G-GAAEGPSAOSTISIRYACRIICDRPPPYTARIYAAGFPAASNNIPIGEBAKKRTDGLM 167
Oy      183  DGLTTNGVLVMMHPRGFTSEESOPGWREISVCGDYVTLTETBSAORQKIVSEETVLOD 242
Db      168  DGLTTNGVLVMMHPRGFTSEESOPGWREISVCGDYVTLTETBSAORQKIVSEETVLOD 227
Oy      243  GSLIDLCGATLLMRTADGLFHPTPOKRIEALRQEOINAAAPQCEVGLNTLAPSPINKEVV 302
Db      228  GSLIDLCGATLLMRTAPAGLLBRAPTLKQLEAQREANAAARQCEVGLSTLAPSPINKEVV 287
Oy      303  BEK-OPMAYLSCGHVHGHNMGHRSSTEANERECPMCRITVGPVYPLMLGCEAGPYVDACP 361
Db      288  PDQOQPMWYVVRGHHVGHYGMWGRBERGRQERCPICRLTVPVYPLMLGCEAGTCLDPGP 347
Oy      362  FTHAFTPCGHVSCSEKSAKYMSOPLPHGTHAFHAAPCPACATOLVGEONCIKLIPOGID 420
Db      348  PSHAAPCGHVSCSEKTAHYMAQPLPHGTHAFHAAPCPGCAWTGHEGCRLLIFQGBLD 406

```

```

RESULT 2
US-11-024-959-511
; Sequence 511, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGOSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 511
; LENGTH: 800
; TYPE: prt
; ORGANISM: Pinus radiata
US-11-024-959-511

```

| | | | | | |
|----|-----------------------|--|----------------|-------|----------------------------------|
| | Query Match | 3.9% | Score 90; | DB 7, | Length 800; |
| | Best Local Similarity | 26.4% | Pred. No. 3.6; | | |
| | Matches | 46; | Conservative | 9; | Mismatches 59; Indels 60; Gaps 6 |
| Oy | 47 | KRPKANGKPSTVHISTPQ-ASKAISCCKG----- | -OHSIS----- | YT | 82 |
| | | : : : | : : : | | |
| Dd | 351 | KBPRTTSLTTTBGLSTPQRAGIAFSSGNLPASSGPPEYVTPPKNGTSNRQTPTTNQGT | | | 410 |
| Oy | 83 | LSRNQTVVV-----EYTHDXTDMFQVGRESTESFDIVTDTTIGSGONTDE---- | | | 128 |
| | | : : : | : : : | | |
| Dd | 411 | LSRPDIIVEIIVPRNSLRPETTSDAKKEANNNGFRVVPSTVSFKSTDVIVSGSNRDESDDKI | | | 470 |
| Oy | 129 | -----AQITQSATISR-----ACRIYDRREPYTARTFAAGF | | | 160 |
| | | : : : | : : : | | |
| Dd | 471 | DSINOKMTGNDKTDTLTARAEQHVASSRDIDNTNTSSVCCDGQOPAPRWGAKKP | | | 524 |

RESULT 3
US-10-995-561-629
; Sequence 629, Application US/10995561
; Publication No. US2005027205A1
; GENERAL INFORMATION:

```

; APPLICANT: CARGILL, Michele et al.
; TITLE OF INVENTION: GENETIC POLYMORPHISMS ASSOCIATED WITH
; TITLE OF INVENTION: CARDIOVASCULAR DISORDERS AND DRUG RESPONSE. METHODS OF
; TITLE OF INVENTION: DETECTION AND USES THEREOF
; FILE REFERENCE: CL001559
; CURRENT APPLICATION NUMBER: US/10/995,561
; CURRENT FILING DATE: 2004-11-24
; NUMBER OF SEQ ID NOS: 85702
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 629
; LENGTH: 984
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-995-561-629

```

[illegible]

```

RESULT 4
US-11-087-099-7273
; Sequence 7273, Application US/11087099
; Publication No. US20060041961A1
; GENERAL INFORMATION:
; APPLICANT: Abad, Mark S. et al.
; TITLE OF INVENTION: Genes and Uses for Plant Improvement
; FILE REFERENCE: 38-21(53450) B EP
; CURRENT APPLICATION NUMBER: US/11/087,099
; CURRENT FILING DATE: 2005-03-22
; NUMBER OF SEQ ID NOS: 12464
; SEQ ID NO 7273
; LENGTH: 452
; TYPE: PR1
; ORGANISM: Nostoc punctiforme
US-11-087-099-7273

```

```

Cy      48  RPKANGVCPVAVHISTPQASKSCGQHSISTLSRN----- 86
Db      51  RKRAVLLGPSNRRLLVDAQENMSRGVYLFSTFGNNILLQGBEHRLLTRRLMYPAFH 110
      Query Match      3.9%; Score 89; DB 7; Length 452;
      Best Local Similarity 22.1%; Pred. No. 2;
      Matches 79; Conservative 32; Mismatches 149; Indels 98; Gaps 14;

```

```

QY 87 -----QTVVVEYTHD-KDTMFOVGRSTESPIDFVVTDTTSGSONTDEAQTOS 134
DB 111 GKAIATVPDIOIIVQDFLKMGERGITSINSSFRQTLTMIATRLPLGSKMSEVE---- 166
QY 135 TISFACRIYCDRNEPYTARIIPAAGPSSKNIFLGEKAATKRP-DGMDGLTTNGVLM 193
DB 167 -----QTSQMFQTLDDSSMAIF-----KMNVEFTLYGRGNARGLVA 204
QY 194 HPRGCTESQPGVMEISVCGVYTLRFRSAQKQKQVSESTNVLDGSLDLCG--- 250
DB 205 FLEAIAQRIBQGNLESK---DVLGLLAVIDEDGNKL--SETOVINEALLLIPAGHET 259
QY 251 -ATILWRADGLHTPTOKHIEALROEINARPOCPVGLNTL-AFPSI-NRKEVEEKOP 307
DB 260 TABLLTWVIFELGNHPEMR--ERLROQLAVGNPNPLSLHLOFQPLTNVAKAEALYR 317
QY 308 WAYSCHGVHYNHGRSDTEANERCEPMCRIVGYPVPLMLCEAGFYVDAGPP-TH 364
DB 318 PVYA-----YNRGVLKDIERYGYRIP-----AGMFTTISPMULTH 351

```

RESULT 5

```

US-11-024-959-471
; Sequence 471, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 471
; LENGTH: 800
; TYPE: PRT
; ORGANISM: Pinus radiata
US-11-024-959-471

```

Query Match 3.9%; Score 89; DB 7; Length 800;

Best Local Similarity 26.4%; Pred. No. 4.5; Matches 46; Conservative 9; Mismatches 59; Indels 60; Gaps 6;

```

QY 47 KRPRKANGVKEPVHVIPTPO-ASKAISCKG-----CHSIS-----YT 82
DB 351 KEKRTTSLTTGSLSTPRAGAFBSKULPASSGPPSVSTPKKNSTSRVQPTTNFOT 410
QY 83 LSNQVTVV-----BYTHDKOTDMFOVGRSTESPIDFVVTDTTSGSONTDE---- 128
DB 411 LSRPDIIVPVIPRNSLAPETTSVKKEMNFGVSVSTSTVDIKSGNSRDESDKI 470
QY 129 -----AQTOSTISR-----ACRIVCDRNEPYTARIIPAAGP 160
DB 471 DSIQKRMGTGNDKTDLNIARAEOHVSSRLDNTNTSSVVCNQPAAKMGIAKAF 524

```

RESULT 6

```

US-11-052-554A-282
; Sequence 282, Application US/11052554A
; Publication No. US2005028866A1
; GENERAL INFORMATION:
; APPLICANT: Sachdeva, et al.
; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
; TITLE OF INVENTION: PROTEINS OF THERAPEUTIC POTENTIAL

```

```

; FILE REFERENCE: 30853/40359A
; CURRENT APPLICATION NUMBER: US/11/052,554A
; CURRENT FILING DATE: 2005-02-07
; PRIOR APPLICATION NUMBER: US 60/589,227
; PRIOR FILING DATE: 2004-07-20
; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
; PRIOR FILING DATE: 2004-02-06
; NUMBER OF SEQ ID NOS: 763
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 282
; LENGTH: 1345
; TYPE: PRT
; ORGANISM: Escherichia coli 0157:H7
US-11-052-554A-282

```

Query Match 3.8%; Score 88; DB 7; Length 1345;

Best Local Similarity 22.6%; Pred. No. 12; Matches 58; Conservative 32; Mismatches 87; Indels 80; Gaps 13;

```

QY 50 KANGVKEPVHVIPTPOASKAISCKQHISVTLSSNOT-----VVVEY 93
DB 90 QANG-----EAVH--TLKSKA-----GTHVTATLSNNNTSDSQPTFVADKTSALVVLQI 139
QY 94 THDKOT-----DMFOVGRSTESPIDFVVTDT-----ISGSONTDEAQTOST 135
DB 140 SKKEITGNGVDATLTAIVKQPD-NEVNNLPVTFSTASGLTLTPGSEVNESGIAQAT 198
QY 136 ISRFACRIYCDRNEPYTARIIPAAGPSSKNIFLGEKAATKRP-DGMDGLTTNGVLM 193
DB 199 LAG-----VAFGEQVYASLANNAGSDNKTVEHFGDTAAKIIETLPVPSIIAGTPONS 253
QY 179 DGHM-----DELTTNGVLMHPRGCTESQPG---VMEISVCGVYTLRFRSAQ 227
DB 254 SCSVITATVVDNNGFPVKGYTVNFTSNAAETNNGQAVTNEQKATVYTT--NTRSSI 311
QY 228 ORGLVSESTNVLDGS 244
DB 312 ESGAPDVTBASLNGS 328

```

RESULT 7

```

US-11-024-959-459
; Sequence 459, Application US/11024959
; Publication No. US20060010516A1
; GENERAL INFORMATION:
; APPLICANT: FORSTER, RICHARD L.
; APPLICANT: CONNETT, MARIE B.
; APPLICANT: EMERSON, SARAH JANE
; APPLICANT: GRIGOR, MURRAY ROBERT
; APPLICANT: HIGGINS, COLLEEN M.
; APPLICANT: LUND, STEVEN TROY
; APPLICANT: MAGUSIN, ANDREAS
; APPLICANT: KODRZYCKI, BOB
; TITLE OF INVENTION: CELL CYCLE GENES AND RELATED METHODS
; FILE REFERENCE: 044463-0360
; CURRENT APPLICATION NUMBER: US/11/024,959
; CURRENT FILING DATE: 2004-12-30
; PRIOR APPLICATION NUMBER: 60/533,036
; PRIOR FILING DATE: 2003-12-30
; NUMBER OF SEQ ID NOS: 782
; SOFTWARE: PatentIn version 3.3
; SEQ ID NO 459
; LENGTH: 1121
; TYPE: PRT
; ORGANISM: Pinus radiata
US-11-024-959-459

```

Query Match 3.8%; Score 87.5; DB 7; Length 1121;

Best Local Similarity 20.9%; Pred. No. 10; Matches 98; Conservative 53; Mismatches 143; Indels 175; Gaps 27;

Db 207 PNGARTPTPTASGLVLPKATPTPLPGHPSPSSSTATSGLASWMSNPMVTSPQAPVA 266
 Qy 71 ISCKGSHISYTSRNOQTVEYTHDKDTOMFOVGRSTESPIDFVYDTIS-----GS 123
 Db 267 V-----GSEVPPPPNATILKRP-----RTPSSSSVVDIQTADSEOLIKRLRPVS 312
 Qy 124 QNTDEA-----QITOST--ISRFACRIVCDRNEPYTARIFAAGFDSKNIFL--GEKA 172
 Db 313 QSIDBATYPPPTLRVPMSTDDLPKTLARAL--NEPVP--VTSIDFHPSQOTFLVGTCKN 367
 Qy 173 AK--WKNPGRHMDGLTTNGVLV-----HPRGGTSESQGVWREISVCGDYTLRE 222
 Db 368 GEITLWE--VGSREKATRSFKIMDNANCSNHLEPAVGVSVSINK-----412
 Qy 223 TRSAQORGLVSEETVNL--QDGSILDLGATLLMTRADGLFHTPPQKHTEALRQ--EIN 278
 Db 413 -----VLSFPGDTLIGIA-----FTKHLVHTYTFQGLD-LRQHLIED 448
 Qy 279 AARPOCPVGLNTLAFPSINKREVEERKQPMAYLSCG-----VHG--YHMGHRS 326
 Db 449 AH-----VGGVNDLAFSHPNKQLC-----VTCGDCKMIKWDATVGRKLYNFEHDA 496
 Qy 327 DTEANEREC-----MCRITYGPVYPLMLGCBAGFYVDAGPPTHAFTPCGHVSEKS 377
 Db 497 PVS--VCPHHKENTOFIFSTAVDGKIKAMLYDHLGSRVDYDAP-----GHSTC--543
 Qy 378 AKWMSQILPHGTHAFHACPCATOLVGE-----QNCITLIRFG 417
 Db 544 ----TMYASADGTRLFS-----CSTKSGESFLVENESGALIKRTYSG 583

RESULT 8

US-11-087-099-7278
 ; Sequence 7278, Application US/11087099
 ; Publication No. US20060041961A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Abad, Mark S. et al.
 ; TITLE OF INVENTION: Genes and Uses for Plant Improvement
 ; FILE REFERENCE: 38-21(53450)B BP
 ; CURRENT APPLICATION NUMBER: US/11/087,099
 ; CURRENT FILING DATE: 2005-03-22
 ; NUMBER OF SEQ ID NOS: 12464
 ; SEQ ID NO 7278
 ; LENGTH: 320
 ; TYPE: PRP
 ; ORGANISM: Glycine max
 ; FEATURE:
 ; NAME/KEY: unsure
 ; LOCATION: (1)..(320)
 ; OTHER INFORMATION: unsure at all Xaa locations
 US-11-087-099-7278

Query Match 3.8%; Score 87; DB 7; Length 320;

Best Local Similarity 22.1%; Pred. No. 1.9;
 Matches 89; Conservative 42; Mismatches 135; Indels 136; Gaps 24;

Qy 28 NGALPNDGRGRKSRFALYKRPKANGVSPVNHVISTPQASK-AISCKGSHISYTSRN 86
 Db 11 NLRLPLPEPSERRPRPLPPTTTIAKPSAGDTIASADLEKLAVALGHGNGGVYKX-RH 69
 Qy 87 QTVVEYTHDKDTOMFOVGRSTESPIDFVYDTISGONTDEAQITOSTISRFA--CRIV 144
 Db 70 KITSATYA-----LKIHSN--ADATRRRAFSSTSLIRATDC---106
 Qy 145 CDRNEPYTARIFAAGFDSKNIFLGEKAAKKNPDG-----HMDGLTTNGVLVHMPRG 197
 Db 107 -----PHVVPFHS-----FENSGVAALIMETDGGTLETALA--TGG 143
 Qy 198 GFTRESQGVWREISVCGDYTLLET--RSAQORGLVSEETVNLQDGSILDLGATLL 254
 Db 144 TFSERLAKYARV-LEGLAVLHAARNIAHRDIKIPANTLVNSEGCV---KIADGVGSKM 198
 Qy 255 WRTADGLFHTPTQGHIEALRQEIWAARPOCPVGLNTLAFPSINKREVEERKQPMAYLSCG 314

Db 199 CRT-----LEA-----CNSYVGFCAVMSPDPRD-----PEAY--GG 227
 Qy 315 HVNGY--HNW-----GHRSDTEANERECRVCGRVYPLW--LGCAGFYVD- 358
 Db 228 NYNGFADWISGLTLFELYVGHFPFLQAGQR-----PWWATLMCAICFSDP 275
 Qy 359 -----ACPTHAFTPCGHVSEK-SAKWMSQILPHGTHAF 393
 Db 276 XSLPETHASRPFHDFVBC-----CLKESGERMTAQLL--THPF 312

RESULT 9

US-11-052-554A-173
 ; Sequence 173, Application US/11052554A
 ; Publication No. US2005028866A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Sachdeva, et al.
 ; TITLE OF INVENTION: COMPUTATIONAL METHOD FOR IDENTIFYING ADHESIN AND ADHESIN-LIKE
 ; FILE REFERENCE: 30853/40359A
 ; CURRENT APPLICATION NUMBER: US/11/052,554A
 ; CURRENT FILING DATE: 2005-02-07
 ; PRIOR APPLICATION NUMBER: US 60/589,227
 ; PRIOR FILING DATE: 2004-07-20
 ; PRIOR APPLICATION NUMBER: IN 173/DEL/2004
 ; PRIOR FILING DATE: 2004-02-06
 ; NUMBER OF SEQ ID NOS: 763
 ; SOFTWARE: PatentIn version 3.3
 ; SEQ ID NO 173
 ; LENGTH: 2105
 ; TYPE: PRP
 ; ORGANISM: Porphyromonas gingivalis W83
 US-11-052-554A-173

Query Match 3.8%; Score 87; DB 7; Length 2105;

Best Local Similarity 22.9%; Pred. No. 27;
 Matches 56; Conservative 35; Mismatches 86; Indels 68; Gaps 17;

Qy 4 PGOEHCAPKKEPVYKGLVYLGYNALPNG-DGRGRKSRFALYKRPKANGVSPVNHV- 61
 Db 714 PGHVAVMISKTGNAGDFVVEE--TPNGINKG--GAFFGL--STANAGARQSWLRE 767
 Qy 62 --ISTPQASKAISCKQH--SISYTSRNOQTVEYTHDKDTOMFOVGRSTESPIDFVVT 117
 Db 768 RTVDLPAGTKYVAFRHNCSDLNYIL-----LDOLQFMGGS-PFTDYTYT 813
 Qy 118 DTISGONTDEAQITOSTISRFACRIVCDRNEPYTARI-FAAG-----FD 161
 Db 814 VYRDGFKIKR--GLTETTFEEDG---VATGNHEVCVEKVTAGVSPKECVNVTVDPVQFN 868
 Qy 162 SKRI--FLGEKAA-KMKNPDGHMDGLTTNGVLVHMPRGFT-----EESQGVWREI 211
 Db 869 PVNLGLSAGVGQVTKLMDAPNG-----PPN-----PNEGTTLSBSFENGIPASWKTI 917
 Qy 212 SVCGD 216
 Db 918 DADGD 922

RESULT 10

US-11-072-512-3891
 ; Sequence 3891, Application US/11072512
 ; Publication No. US2006002945A1
 ; GENERAL INFORMATION:
 ; APPLICANT: ISOGAI, TAKAO
 ; APPLICANT: SUGIYAMA, TOMOYASU
 ; APPLICANT: OTSUKI, TETSUJI
 ; APPLICANT: WAKAMATSU, AI
 ; APPLICANT: SATO, HIROYUKI
 ; APPLICANT: ISHII, SHIZUKO
 ; APPLICANT: YAMAMOTO, JUN-ICHI
 ; APPLICANT: ISONO, YUUKO

APPLICANT: HIO, YURI
APPLICANT: OTSUKA, KAORU
APPLICANT: NAGAI, KEIICHI
APPLICANT: IRIE, RYOTARO
APPLICANT: TAMECHIKA, ICHIRO
APPLICANT: SEKI, NAOHICO
APPLICANT: YOSHIKAWA, TSUTOMU
APPLICANT: OTSUKA, MOTOKYU
APPLICANT: NAGAHARI, KENJI
APPLICANT: MASUHO, YASUHIKO
TITLE OF INVENTION: Novel full length cDNA
FILE REFERENCE: 084335-0191
CURRENT APPLICATION NUMBER: US/11/072,512
CURRENT FILING DATE: 2005-03-07
PRIOR APPLICATION NUMBER: US 60/350,978
PRIOR FILING DATE: 2002-01-25
PRIOR APPLICATION NUMBER: JP 2001-379298
PRIOR FILING DATE: 2001-11-05
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 3891
LENGTH: 266
TYPE: PRT
ORGANISM: Homo sapiens
US-11-072-512-3891

Query Match 3.8%; Score 86.5; DB 7; Length 266;
Best Local Similarity 23.1%; Pred. No. 1.6;
Matches 45; Conservative 20; Mismatches 81; Indels 49; Gaps 9;

QY 225 SAOQKGLVSEENVLQDGS---SLIDCGATLLM-RTADGLFHTPTQKHTEALQGEINAA 280
DB 29 SGQKGGALMSGLCTIGKHLSTLNLGPKVVCDDKHPQGL---DNHFFLLFLLSGA 84
QY 281 RPOCPVGLATLAFPSINRKEVEE---KQPMAYLSCGHVGHYHNGHRSDEANERECPM 337
DB 85 PKSTWTALMTSGYKRETFQESVMTAGSEPAAYLKRAHHNN-HYTWISICKKXNPRRCVS 143
QY 338 CRTVGPVPLMLGCEAGPYVDAGPPTAFTPCCGVCSKSAKYWSQIPLPHGTHAFHA- 396
DB 144 -----GAAKRIRHKGADPS-----IPLPHQSASFICD 171
QY 397 -CPFCATQVGEQNC 410
DB 172 RCPHC---LSGPAC 183

RESULT 11
US-11-149-003-22

Sequence 22, Application US/11149003
Publication No. US20060014277A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Scoville, John
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: Novel Human Kiehn-like Proteins and Polynucleotides Encoding the
FILE REFERENCE: LEX-0360-USA
CURRENT APPLICATION NUMBER: US/11/149,003
CURRENT FILING DATE: 2005-06-09
PRIOR APPLICATION NUMBER: US/10/189,971
PRIOR FILING DATE: 2002-07-03
PRIOR APPLICATION NUMBER: US 60/302,949
PRIOR FILING DATE: 2001-07-03
PRIOR APPLICATION NUMBER: US 60/315,634
PRIOR FILING DATE: 2001-08-29
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 22
LENGTH: 759
TYPE: PRT
ORGANISM: homo sapiens
US-11-149-003-22

Query Match 3.8%; Score 86; DB 7; Length 759;
Best Local Similarity 20.5%; Pred. No. 7.9; Indels 100; Gaps 17;
Matches 64; Conservative 28; Mismatches 120; Indels 100; Gaps 17;

QY 166 IFLEKKAQKKNP-----DGHMDGL-----TTNGVLVHMPRGF-----T 200
DB 463 VILGDMAVRLDQGAVTVDHPPALPQLGRLLYVELRGHTVILHQPGLVLMDSQSV 522
QY 201 EESQPGTW--REISVCGDYTTLAEFTSAQQRGLVSEENVLQDGSILIDCGATLLMRTA 258
DB 523 EVSVPGSYQGRTCGLCGNFNGFAQDDLOQGEGLLPSSEAFAFGNS-----WQVS 570
QY 259 DGLF---HTPTQKHTEALROEINARPPQCVGINTLAFPSINRKEVEEKOP-----WA 309
DB 571 EGLMPGRCAGRBVDPCRAAGYRAREANARCGVLKSPFSRCHAVPPEPPAACVVD 630
QY 310 YLSCGHVGHYHNGHRSDE---EANERECPMCRVTGPPVPLM-----LGC--EAGFPV 357
DB 631 LCAQGP-----GSSADACLCDALAEVYASHCRQAG--VPTWAGPTLCVCGCPLERGFV 682
QY 358 D-AGPP-----THAFTPCGVCSKSAKYWSQIPLPHGTHAFHA---A 396
DB 683 DECGPFCPTCFNQHIPLGELAAHCVRPVPGC-----CCPAGLVHBAHCIPPEA 733
QY 397 CPFCATQVGEQ 408
DB 734 CP--QVLLTNDQ 743

RESULT 12
US-11-149-003-6

Sequence 6, Application US/11149003
Publication No. US20060014277A1
GENERAL INFORMATION:
APPLICANT: Walke, D. Wade
APPLICANT: Scoville, John
APPLICANT: Turner, C. Alexander Jr.
TITLE OF INVENTION: Novel Human Kiehn-like Proteins and Polynucleotides Encoding the
FILE REFERENCE: LEX-0360-USA
CURRENT APPLICATION NUMBER: US/11/149,003
CURRENT FILING DATE: 2005-06-09
PRIOR APPLICATION NUMBER: US/10/189,971
PRIOR FILING DATE: 2002-07-03
PRIOR APPLICATION NUMBER: US 60/302,949
PRIOR FILING DATE: 2001-07-03
PRIOR APPLICATION NUMBER: US 60/315,634
PRIOR FILING DATE: 2001-08-29
NUMBER OF SEQ ID NOS: 25
SOFTWARE: FastSeq for Windows Version 4.0
SEQ ID NO 6
LENGTH: 1057
TYPE: PRT
ORGANISM: homo sapiens
US-11-149-003-6

Query Match 3.8%; Score 86; DB 7; Length 1057;
Best Local Similarity 20.5%; Pred. No. 13;
Matches 64; Conservative 28; Mismatches 120; Indels 100; Gaps 17;

QY 166 IFLEKKAQKKNP-----DGHMDGL-----TTNGVLVHMPRGF-----T 200
DB 761 VILGDMAVRLDQGAVTVDHPPALPQLGRLLYVELRGHTVILHQPGLVLMDSQSV 820
QY 201 EESQPGTW--REISVCGDYTTLAEFTSAQQRGLVSEENVLQDGSILIDCGATLLMRTA 258
DB 821 EVSVPGSYQGRTCGLCGNFNGFAQDDLOQGEGLLPSSEAFAFGNS-----WQVS 868
QY 259 DGLF---HTPTQKHTEALROEINARPPQCVGINTLAFPSINRKEVEEKOP-----WA 309
DB 869 EGLMPGRCAGRBVDPCRAAGYRAREANARCGVLKSPFSRCHAVPPEPPAACVVD 928

[illegible]

```

RESULT 13
US-11-149-003-18
: Sequence 18, Application US/11149003
: Publication No. US20060014277A1
: GENERAL INFORMATION:
: APPLICANT: Walke, D. Wade
: APPLICANT: Scoville, John
: APPLICANT: Turner, C. Alexander Jr.
: TITLE OF INVENTION: Novel Human Kiflin-like Proteins and Polynucleotides Encoding the
: TITLE OF INVENTION: Same
: FILE REFERENCE: LEX-0360-USA
: CURRENT APPLICATION NUMBER: US/11/149, 003
: CURRENT FILING DATE: 2005-06-09
: PRIOR APPLICATION NUMBER: US/10/189,971
: PRIOR FILING DATE: 2002-07-03
: PRIOR APPLICATION NUMBER: US 60/302,949
: PRIOR FILING DATE: 2001-07-03
: PRIOR APPLICATION NUMBER: US 60/315,634
: PRIOR FILING DATE: 2001-08-29
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ ID NO 18
: LENGTH: 1192
: TYPE: PRT
: ORGANISM: homo sapiens
: US-11-149-003-18

```

| | | | | |
|-----------------------|-------|---|-------------------------------------|--------------------------|
| Query Match | 3.8% | Score 86; | DB 7; | Length 1192; |
| Best Local Similarity | 20.5% | Pred. No. 15; | | |
| Matches | 64; | Conservative | 28; | Mismatches 120; |
| | | | Indels 100; | Gaps 17; |
| QY | 166 | IFLGEKAKXWKNP----- | DGHMDGL----- | TTNGVLVMEPRGAF-----T 200 |
| | | | | |
| | | | | |
| Db | 896 | VLLDDMAVRLLDGAVTVTDGHPVALPFLQEPPLLVELNGHTYILHAQGLQVLMQDSQV | | 955 |
| QY | 201 | ESQSGVW--REISVCGDYTLIRETRSAQQRKGLVESETNVLQDSSLIDLCGATILMRTA | | 258 |
| | | | | |
| | | | | |
| Db | 956 | EVSVPGSYQGTCTGICGNFNGPAQDDLGQPBGLLPSAAFGNS----- | WQVS | 1003 |
| QY | 259 | DGLF---HTTPQKHLEALROEINARPCQPGVGLNTLAPRINRKEVEBEKQ----- | WA | 309 |
| | | | | |
| Db | 1004 | KALMPGRCSAGREYVDPCLAAAGYRARREANARCGVLKSSPFSRCHAVVPPEPFACAYD | | 1063 |
| QY | 310 | YLSCGHVGHYNNWGRISDT--EANERECPCMRCTGPPVPLW----- | LGC--EAGEYV | 357 |
| | | | | |
| | | | | |
| Db | 1064 | LCAGSP-----GSSADACLDALAEAAASHRQNG-VTPTRRGPTLCVSGCPLERGTF | | 1115 |
| QY | 358 | D-ACGP----- | THAFTPCGHVCSKSAKYSQIDLPHTHAFAH----- | A 396 |
| | | | | |
| | | | | |
| Db | 1116 | DECGPPCPRTCFNQHIFLGEILAAHCVRPCVPC----- | QCPAGLVEHAHCIPREA | 1166 |
| QY | 397 | CPFCATQLVGEQ | 408 | |
| | | | | |
| | | | | |
| Db | 1167 | CP--QVLLTQDQ | 1176 | |

RESULT 14
US-11-149-003-20
; Sequence 20, Application US/11149003
; Publication No. US20060014277A1

```

: GENERAL INFORMATION:
: APPLICANT: Walke, D. Wade
: APPLICANT: Scoville, John
: APPLICANT: Turner, C. Alexander Jr.
: TITLE OF INVENTION: Novel Human Kiehn-like Proteins and Polynucleotides Encoding the
: TITLE OF INVENTION: Same
: FILE REFERENCE: LEX-0360-USA
: CURRENT APPLICATION NUMBER: US/11/149,003
: CURRENT FILING DATE: 2005-06-09
: PRIOR APPLICATION NUMBER: US/10/189,971
: PRIOR FILING DATE: 2002-07-03
: PRIOR APPLICATION NUMBER: US 60/302,949
: PRIOR FILING DATE: 2001-07-03
: PRIOR APPLICATION NUMBER: US 60/315,634
: PRIOR FILING DATE: 2001-08-29
: NUMBER OF SEQ. ID NOS: 25
: SOFTWARE: FastSeq for Windows Version 4.0
: SEQ. ID NO: 20
: LENGTH: 1207
: TYPE: PRT
: ORGANISM: homo sapiens
: US-11-149-003-20

Query Match          3.84; Score 86; DB 7; Length 1207;
Best Local Similarity 20.5%; Pred. No. 15;
Matches 64; Conservative 28; Mismatches 100; Gaps 17;

```

| | | |
|----|---|---|
| | Query Match | 3.8%; Score 86; DB 7; Length 1207; |
| | Best Local Similarity | 20.5%; Pred. No. 15; |
| | Matches | 64; Conservative 28; Mismatches 120; Indels 100; Gaps 17; |
| Qy | 166 IFLGEKAKAKMKNP-----DGHMDGL-----TTNGVLVNHPRGGF-----T | 200 |
| Db | 911 VLLGDMAVRLLDGAVTVVDGHPRVALPFLQEPLLYVELRGHTVILHAQPGQLVLMDSQS | 970 |
| Qy | 201 EESDPGW--REISVGEDVTTLRRTSAQRGKLVESETNVLDGSLDLDCATLLWRTA | 258 |
| Db | 971 EVSPGSYGQGTCCGLCANFNFGPADDDLOGEPGLLPSEAAFGNS-----MQVS | 1018 |
| Qy | 259 DGLF---HTPPQKHIALROEINARPOCPYGLNTLLAPSIINKEVVEBKOP-----WA | 309 |
| Db | 1019 EGLMPGRPCSAKRVDVPCRAAGYARRAEANARCGLVASSPPSRCHAVPEPPFPFAACYTD | 1078 |
| Qy | 310 YLSGHHVGYHNMGHRSDT--EANECEPCWCTGVGPVPLM-----LGC--EAGEYV | 357 |
| Db | 1079 LCACGP-----GSSADACLCDLLEVAISHCRDAG--VTPWRGPTLCVVCSPERGVVF | 1130 |
| Qy | 358 D-AGRP-----THAFTPCGHVCSSEKSACTWSQIPLPHGTAFFA----A | 396 |
| Db | 1131 DECPCPCPRTCFNQHIDPLGETLAHCVRVPVGC-----QCPSGLVEHEACPPEEA | 1181 |
| Qy | 397 CPFCAQTQLVBEQ | 408 |
| Db | 1182 CP--QVILLTGQQ | 1191 |

```

RESULT 15
US-11-149-003-16
: Sequence 16. Application US/11149003
: Publication No. US20060014277A1
: GENERAL INFORMATION:
: APPLICANT: Walke, D. Wade
: APPLICANT: Scoville, John
: APPLICANT: Turner, C. Alexander Jr.
: TITLE OF INVENTION: Novel Human K14-like Proteins and Polynucleotides Encoding the
: TITLE OF INVENTION: Same
: FILE REFERENCE: LEX-0360-USA
: CURRENT APPLICATION NUMBER: US/11/149,003
: CURRENT FILING DATE: 2005-06-09
: PRIOR APPLICATION NUMBER: US/10/189,971
: PRIOR FILING DATE: 2002-07-03
: PRIOR APPLICATION NUMBER: US 60/302,949
: PRIOR FILING DATE: 2001-07-03
: PRIOR APPLICATION NUMBER: US 60/315,634
: PRIOR FILING DATE: 2001-08-29
: NUMBER OF SEQ ID NOS: 25
: SOFTWARE: FastSeq for Windows Version 4.0

```


SEQ ID NO 16
LENGTH: 1251
TYPE: PRT
ORGANISM: homo sapiens
US-11-149-003-16

Query Match 3.8%; Score 86; DB 7; Length 1251;
Best Local Similarity 20.5%; Pred. No. 16;
Matches 64; Conservative 28; Mismatches 120; Indels 100; Gaps 17;

| | | | |
|----|------|--|------|
| QY | 166 | IFLGEKAKAKWNP-----DGHMDGL-----TTNGVLVWHPRGGF-----T | 200 |
| DB | 955 | VLLGDMAVRLIQDGAVTVDGHVPLPFLQBPPLVYELRGHTVILHAQPGLOVMDGQSOV | 1014 |
| QY | 201 | EESSPGTW--REISVCGDVYTLRETRSAQQRKLYESETNVLQDGLIDLGCATILWRTA | 258 |
| DB | 1015 | EVSVPGSYGRTCGICGNFNGPAQDDLOGPBGILLPSEAFGNS-----MQVS | 1062 |
| QY | 259 | DGLF---HTPTQKHIEALQRIINARPCQPVGLNTLAFPSINRKVVEKOP-----WA | 309 |
| DB | 1063 | EGLMFGRPCSGAGREVDPCRAAGYRAAREANARCGVLKSSPFSRCHAVVPPEPFAACYID | 1122 |
| QY | 310 | YLSGCVHVGYNWGHRSPT--KANRECEPMCRVTGVPYPLM-----LGC--EAGFYV | 357 |
| DB | 1123 | LCACGP-----GSSADACLCDALEAVASHCRQAG--VTPWRGPPLCVGCPLERGFVF | 1174 |
| QY | 358 | D-AGPP-----THAFTPCGHVCESEKSAKYSQIPLPHGTHAFHA---A | 396 |
| DB | 1175 | DECBPPCPCRTCFNOHPIPLGELAAHCVRPCVPGC-----QCPAGLVEHEAHCIPEEA | 1225 |
| QY | 397 | CPFCATOLVGEQ | 408 |
| DB | 1226 | CP--QVLITDQ | 1235 |

Search completed: March 1, 2006, 21:04:59
Job time : 21 secs

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2006, 21:01:01 ; Search time 167 Seconds
(without alignments)
1050.828 Million cell updates/sec

Title: US-10-041-030-4

Sequence: 1 MESPGEHEKAPKPEPVKYG.....ATQVGEQNCIKILFQGPID 420

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 1867569 seqs, 417829326 residues

Total number of hits satisfying chosen parameters: 1867569

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database: Published Applications AA Main:
1: /cgn2_6/ptodaca/1/pubpaa/US07_PUBCOMB.pep:*
2: /cgn2_6/ptodaca/1/pubpaa/US08_PUBCOMB.pep:*
3: /cgn2_6/ptodaca/1/pubpaa/US09_PUBCOMB.pep:*
4: /cgn2_6/ptodaca/1/pubpaa/US10A_PUBCOMB.pep:*
5: /cgn2_6/ptodaca/1/pubpaa/US10B_PUBCOMB.pep:*
6: /cgn2_6/ptodaca/1/pubpaa/US11_PUBCOMB.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|---------------------|--------------------|
| 1 | 2290 | 100.0 | 420 | US-10-041-030-4 | Sequence 4, Appl |
| 2 | 2290 | 100.0 | 420 | US-10-197-666A-136 | Sequence 136, App |
| 3 | 2283 | 99.7 | 420 | US-09-843-905A-8 | Sequence 8, Appl |
| 4 | 2283 | 99.7 | 420 | US-10-317-250-8 | Sequence 8, Appl |
| 5 | 2283 | 99.7 | 420 | US-10-258-703-8 | Sequence 8, Appl |
| 6 | 2175.5 | 95.0 | 419 | US-09-843-905A-6 | Sequence 6, Appl |
| 7 | 2175.5 | 95.0 | 419 | US-10-317-250-6 | Sequence 6, Appl |
| 8 | 2175.5 | 95.0 | 419 | US-10-258-703-6 | Sequence 6, Appl |
| 9 | 2111.5 | 92.2 | 419 | US-10-197-666A-134 | Sequence 134, App |
| 10 | 1917 | 83.7 | 418 | US-09-843-905A-4 | Sequence 4, Appl |
| 11 | 1917 | 83.7 | 418 | US-10-317-250-4 | Sequence 4, Appl |
| 12 | 1917 | 83.7 | 418 | US-10-443-108-10 | Sequence 10, Appl |
| 13 | 1917 | 83.7 | 418 | US-10-258-703-4 | Sequence 4, Appl |
| 14 | 1913 | 83.5 | 418 | US-10-041-030-2 | Sequence 2, Appl |
| 15 | 1912 | 83.5 | 418 | US-09-843-905A-2 | Sequence 2, Appl |
| 16 | 1912 | 83.5 | 418 | US-10-317-250-2 | Sequence 2, Appl |
| 17 | 1912 | 83.5 | 418 | US-10-258-703-2 | Sequence 2, Appl |
| 18 | 1642 | 71.7 | 445 | US-09-843-905A-12 | Sequence 12, Appl |
| 19 | 1642 | 71.7 | 445 | US-10-317-250-12 | Sequence 12, Appl |
| 20 | 1642 | 71.7 | 445 | US-10-258-703-12 | Sequence 12, Appl |
| 21 | 1482.5 | 64.7 | 406 | US-10-104-047-3928 | Sequence 3928, App |
| 22 | 1302 | 56.9 | 424 | US-09-843-905A-13 | Sequence 13, Appl |
| 23 | 1302 | 56.9 | 424 | US-10-317-250-13 | Sequence 13, Appl |
| 24 | 1302 | 56.9 | 424 | US-10-258-703-13 | Sequence 13, Appl |
| 25 | 1302 | 56.9 | 424 | US-11-097-143-10851 | Sequence 10851, A |
| 26 | 1235 | 53.9 | 455 | US-09-843-905A-14 | Sequence 14, Appl |
| 27 | 1235 | 53.9 | 455 | US-10-317-250-14 | Sequence 14, Appl |

| | | | | | | |
|----|------|------|------|---|---------------------|-------------------|
| 28 | 1235 | 53.9 | 455 | 4 | US-10-258-703-14 | Sequence 14, Appl |
| 29 | 1077 | 47.0 | 241 | 5 | US-10-450-763-34853 | Sequence 34853, A |
| 30 | 974 | 42.5 | 171 | 4 | US-10-029-386-32174 | Sequence 32174, A |
| 31 | 903 | 39.4 | 458 | 3 | US-09-843-905A-15 | Sequence 15, Appl |
| 32 | 903 | 39.4 | 458 | 4 | US-10-317-250-15 | Sequence 15, Appl |
| 33 | 903 | 39.4 | 458 | 4 | US-10-258-703-15 | Sequence 15, Appl |
| 34 | 775 | 33.8 | 155 | 3 | US-09-867-550-1948 | Sequence 1948, Ap |
| 35 | 678 | 29.6 | 125 | 3 | US-09-867-550-1720 | Sequence 1720, Ap |
| 36 | 343 | 15.0 | 528 | 5 | US-10-450-763-34852 | Sequence 34852, A |
| 37 | 154 | 6.7 | 26 | 4 | US-10-041-030-38 | Sequence 38, Appl |
| 38 | 113 | 4.9 | 818 | 5 | US-10-732-923-14186 | Sequence 14186, A |
| 39 | 113 | 4.9 | 818 | 5 | US-10-732-923-14187 | Sequence 14187, A |
| 40 | 109 | 4.8 | 1200 | 5 | US-10-450-763-36074 | Sequence 36074, A |
| 41 | 109 | 4.8 | 1379 | 5 | US-10-450-763-36078 | Sequence 36075, A |
| 42 | 104 | 4.5 | 20 | 4 | US-10-041-030-26 | Sequence 26, Appl |
| 43 | 100 | 4.4 | 19 | 4 | US-10-041-080-31 | Sequence 31, Appl |
| 44 | 100 | 4.4 | 1114 | 5 | US-10-722-923-14185 | Sequence 14185, A |
| 45 | 100 | 4.4 | 2042 | 4 | US-10-192-584-6 | Sequence 6, Appl |

ALIGNMENTS

RESULT 1
US-10-041-030-4
Sequence 4, Application US/10041030
Publication No. US20020150934A1
GENERAL INFORMATION:
APPLICANT: Powers, Scott
APPLICANT: Ma, David
APPLICANT: Xiang, Phil
APPLICANT: Peng, Yue
APPLICANT: Tularik Inc.
TITLE OF INVENTION: Diagnosis and Treatment of Cancer Using Mammalian
FILE REFERENCE: 018781-006810US
CURRENT APPLICATION NUMBER: US/10/041, 030
CURRENT FILING DATE: 2001-12-28
PRIOR APPLICATION NUMBER: US 60/259, 502
PRIOR FILING DATE: 2001-01-02
NUMBER OF SEQ ID NOS: 42
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 4
LENGTH: 420
TYPE: PRT
ORGANISM: Homo sapiens
FEATURES:
OTHER INFORMATION: human pellino 2
US-10-041-030-4
Query Match 100.0%; Score 2290; DB 4; Length 420;
Best Local Similarity 100.0%; Pred. No. 1.6e-216;
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
1 MESPGEHEKAPKPEPVKYGELVLYGNGALPMGDRGRKSRALYKPPKANGVKSTVH 60
1 MESPGEHEKAPKPEPVKYGELVLYGNGALPMGDRGRKSRALYKPPKANGVKSTVH 60
61 VITPQASKAISCKGSHSYTLRSNQTIVVEYTHKQDMDPMFGVSTESPIDFVVTDTI 120
61 VITPQASKAISCKGSHSYTLRSNQTIVVEYTHKQDMDPMFGVSTESPIDFVVTDTI 120
61 VITPQASKAISCKGSHSYTLRSNQTIVVEYTHKQDMDPMFGVSTESPIDFVVTDTI 120
121 SGGQNTDEAQITOSTISRFACRIVCDRNEPYTARIPAGFSSKNIPLGSKAKWKNPDG 180
121 SGGQNTDEAQITOSTISRFACRIVCDRNEPYTARIPAGFSSKNIPLGSKAKWKNPDG 180
121 SGGQNTDEAQITOSTISRFACRIVCDRNEPYTARIPAGFSSKNIPLGSKAKWKNPDG 180
181 HMGITLNGVLVHMPRGFTRESQPGVWREISVCGGVYTLRETRSAQQRKLVESNTVL 240
181 HMGITLNGVLVHMPRGFTRESQPGVWREISVCGGVYTLRETRSAQQRKLVESNTVL 240
241 QDSGLIDLCATILMTADGLFTTPQOKITEALROEINARPOCPGINTLAFPSINRKE 300
241 QDSGLIDLCATILMTADGLFTTPQOKITEALROEINARPOCPGINTLAFPSINRKE 300

QY 301 VVEEKOPWATLSCGHVGHYNNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
Db 301 VVEEKOPWATLSCGHVGHYNNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
QY 361 PPTHAFPCGHVCGHVSSEKSAKYSOIPLPHGTHAFHAACPCATOLVGEONCICKLIFQGPID 420
Db 361 PPTHAFPCGHVCGHVSSEKSAKYSOIPLPHGTHAFHAACPCATOLVGEONCICKLIFQGPID 420

RESULT 2

US-10-197-666A-136
; Sequence 136, Application US/10197666A
; Publication No. US2003092037A1
; GENERAL INFORMATION:
; APPLICANT: ASAMI KASEI KABUSIKI KAISYA
; TITLE OF INVENTION: ELK1 phosphorylation related gene
; FILE REFERENCE: PH-1548US
; CURRENT APPLICATION NUMBER: US/10/197,666A
; PRIOR FILING DATE: 2002-11-18
; PRIOR APPLICATION NUMBER: JP 2001-218204
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: JP 2001-263450
; PRIOR FILING DATE: 2001-08-31
; PRIOR APPLICATION NUMBER: JP 2002-012176
; PRIOR FILING DATE: 2002-01-21
; PRIOR APPLICATION NUMBER: US 60/305,884
; PRIOR FILING DATE: 2001-07-18
; PRIOR APPLICATION NUMBER: US 60/316,304
; PRIOR FILING DATE: 2001-09-04
; PRIOR APPLICATION NUMBER: US 60/350,027
; PRIOR FILING DATE: 2002-01-23
; NUMBER OF SEQ ID NOS: 156
; SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 136
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-197-666A-136

Query Match 100.0%; Score 2290; DB 4; Length 420;
Best Local Similarity 100.0%; Pred. No. 1.6e-216;
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSPGQEHCAPKKEPVKYGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTVH 60
Db 1 MFSPGQEHCAPKKEPVKYGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTVH 60
QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVEYTHDKDTMFOVGRSTESPIDFVVTDIT 120
Db 61 VISTPQASKAISCKGHSISYTLSSRNQTVVEYTHDKDTMFOVGRSTESPIDFVVTDIT 120
QY 121 SSGQNTDEAOITOSTISRPAFCRIYCDRNEPYTARIFAAGFDSKNIIFLGKAKAKMKNPDG 180
Db 121 SSGQNTDEAOITOSTISRPAFCRIYCDRNEPYTARIFAAGFDSKNIIFLGKAKAKMKNPDG 180
QY 181 HMDGLTTNGVLVNHPRGFTBESOPGVWREISVCGDVYTLRETRSAOORCKLVESENNVL 240
Db 181 HMDGLTTNGVLVNHPRGFTBESOPGVWREISVCGDVYTLRETRSAOORCKLVESENNVL 240
QY 241 QDSGLIDLCGATILMTADGLFHTPTQKIEALROEINARPOCPVGLNTLAFPSINRKE 300
Db 241 QDSGLIDLCGATILMTADGLFHTPTQKIEALROEINARPOCPVGLNTLAFPSINRKE 300
QY 301 VVEEKOPWATLSCGHVGHYNNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
Db 301 VVEEKOPWATLSCGHVGHYNNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
QY 361 PPTHAFPCGHVCGHVSSEKSAKYSOIPLPHGTHAFHAACPCATOLVGEONCICKLIFQGPID 420
Db 361 PPTHAFPCGHVCGHVSSEKSAKYSOIPLPHGTHAFHAACPCATOLVGEONCICKLIFQGPID 420

RESULT 3

US-09-843-905A-8
; Sequence 8, Application US/09843905A
; Patent No. US2002016683A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; PRIOR FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo sapiens
US-09-843-905A-8

Query Match 99.7%; Score 2283; DB 3; Length 420;
Best Local Similarity 99.8%; Pred. No. 7.7e-216;
Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFSPGQEHCAPKKEPVKYGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTVH 60
Db 1 MFSPGQEHCAPKKEPVKYGELVVLGYNGALPNDGRGRKSRPALYKRPKANGVKPSTVH 60
QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVEYTHDKDTMFOVGRSTESPIDFVVTDIT 120
Db 61 VISTPQASKAISCKGHSISYTLSSRNQTVVEYTHDKDTMFOVGRSTESPIDFVVTDIT 120
QY 121 SSGQNTDEAOITOSTISRPAFCRIYCDRNEPYTARIFAAGFDSKNIIFLGKAKAKMKNPDG 180
Db 121 SSGQNTDEAOITOSTISRPAFCRIYCDRNEPYTARIFAAGFDSKNIIFLGKAKAKMKNPDG 180
QY 181 HMDGLTTNGVLVNHPRGFTBESOPGVWREISVCGDVYTLRETRSAOORCKLVESENNVL 240
Db 181 HMDGLTTNGVLVNHPRGFTBESOPGVWREISVCGDVYTLRETRSAOORCKLVESENNVL 240
QY 241 QDSGLIDLCGATILMTADGLFHTPTQKIEALROEINARPOCPVGLNTLAFPSINRKE 300
Db 241 QDSGLIDLCGATILMTADGLFHTPTQKIEALROEINARPOCPVGLNTLAFPSINRKE 300
QY 301 VVEEKOPWATLSCGHVGHYNNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
Db 301 VVEEKOPWATLSCGHVGHYNNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
QY 361 PPTHAFPCGHVCGHVSSEKSAKYSOIPLPHGTHAFHAACPCATOLVGEONCICKLIFQGPID 420
Db 361 PPTHAFPCGHVCGHVSSEKSAKYSOIPLPHGTHAFHAACPCATOLVGEONCICKLIFQGPID 420

RESULT 4

US-10-317-250-8
; Sequence 8, Application US/10317250
; Publication No. US20030165945A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; APPLICANT: Li, Xiaoxia
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-B
; CURRENT APPLICATION NUMBER: US/10/317,250
; PRIOR FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 8
; LENGTH: 420
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-317-250-8

Query Match 99.7%; Score 2283; DB 4; Length 420;
 Best Local Similarity 99.8%; Pred. No. 7.7e-216;
 Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFSPGQEHCAPNKEPVKYGELVVLGYNGALPNGDRGRKRSFALYKRPANGVKSTVH 60
 DB 1 MFSPGQEHCAPNKEPVKYGELVVLGYNGALPNGDRGRKRSFALYKRPANGVKSTVH 60
 QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVEYTHDKOTDMFOVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQASKAISCKGHSISYTLSSRNQTVVEYTHDKOTDMFOVGRSTESPIDFVVTDTI 120
 QY 121 SSGQNTDEAQTOSTISRFACRIYCDNREPYTARIIPAAGDSSKNITLGEKAKMKNPDG 180
 DB 121 SSGQNTDEAQTOSTISRFACRIYCDNREPYTARIIPAAGDSSKNITLGEKAKMKNPDG 180
 QY 181 HMDGLTNGVLVNHMPRGFTESQPGVWREISVCGDVYTLRETRSAQQRGLVSESTNVL 240
 DB 181 HMDGLTNGVLVNHMPRGFTESQPGVWREISVCGDVYTLRETRSAQQRGLVSESTNVL 240
 QY 241 QDQSLIDLCGATLLMRTADGLFHTPTQKHIALROBINARPOCPVGLNTLAPPSINRKE 300
 DB 241 QDQSLIDLCGATLLMRTADGLFHTPTQKHIALROBINARPOCPVGLNTLAPPSINRKE 300
 QY 301 VVEBKQPMAYLSCGHVGHYHNMGRSDTEANERECPCMCRTVGPVPLMLGCEAGFYVDAG 360
 DB 301 VVEBKQPMAYLSCGHVGHYHNMGRSDTEANERECPCMCRTVGPVPLMLGCEAGFYVDAG 360
 QY 361 PPTHAFTPCGHVSSEKSAKXWSQIPLPHGTHAFHAACPFCATQLVGBQNCIKLIFQGPID 420
 DB 361 PPTHAFTPCGHVSSEKSAKXWSQIPLPHGTHAFHAACPFCATQLVGBQNCIKLIFQGPID 420

RESULT 5

US-10-258-703-8
 ; Sequence 8, Application US/10258703
 ; Publication No. US20040034199A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bird, Timothy A.
 ; APPLICANT: Cosman, David J.
 ; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
 ; FILE REFERENCE: 2990-US
 ; CURRENT APPLICATION NUMBER: US/10/258,703
 ; CURRENT FILING DATE: 2002-10-24
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 8
 ; LENGTH: 420
 ; TYPE: PRT
 ; ORGANISM: Homo sapiens
 ; US-10-258-703-8

Query Match 99.7%; Score 2283; DB 4; Length 420;
 Best Local Similarity 99.8%; Pred. No. 7.7e-216;
 Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MFSPGQEHCAPNKEPVKYGELVVLGYNGALPNGDRGRKRSFALYKRPANGVKSTVH 60
 DB 1 MFSPGQEHCAPNKEPVKYGELVVLGYNGALPNGDRGRKRSFALYKRPANGVKSTVH 60
 QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVEYTHDKOTDMFOVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQASKAISCKGHSISYTLSSRNQTVVEYTHDKOTDMFOVGRSTESPIDFVVTDTI 120
 QY 121 SSGQNTDEAQTOSTISRFACRIYCDNREPYTARIIPAAGDSSKNITLGEKAKMKNPDG 180
 DB 121 SSGQNTDEAQTOSTISRFACRIYCDNREPYTARIIPAAGDSSKNITLGEKAKMKNPDG 180
 QY 181 HMDGLTNGVLVNHMPRGFTESQPGVWREISVCGDVYTLRETRSAQQRGLVSESTNVL 240
 DB 181 HMDGLTNGVLVNHMPRGFTESQPGVWREISVCGDVYTLRETRSAQQRGLVSESTNVL 240

QY 241 QDQSLIDLCGATLLMRTADGLFHTPTQKHIALROBINARPOCPVGLNTLAPPSINRKE 300
 DB 241 QDQSLIDLCGATLLMRTADGLFHTPTQKHIALROBINARPOCPVGLNTLAPPSINRKE 300
 QY 301 VVEBKQPMAYLSCGHVGHYHNMGRSDTEANERECPCMCRTVGPVPLMLGCEAGFYVDAG 360
 DB 301 VVEBKQPMAYLSCGHVGHYHNMGRSDTEANERECPCMCRTVGPVPLMLGCEAGFYVDAG 360
 QY 361 PPTHAFTPCGHVSSEKSAKXWSQIPLPHGTHAFHAACPFCATQLVGBQNCIKLIFQGPID 420
 DB 361 PPTHAFTPCGHVSSEKSAKXWSQIPLPHGTHAFHAACPFCATQLVGBQNCIKLIFQGPID 420

RESULT 6

US-09-843-905A-6
 ; Sequence 6, Application US/09843905A
 ; Patent No. US20020168683A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bird, Timothy A.
 ; APPLICANT: Cosman, David J.
 ; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
 ; FILE REFERENCE: 2990-A
 ; CURRENT APPLICATION NUMBER: US/09/843,905A
 ; CURRENT FILING DATE: 2001-04-27
 ; PRIOR APPLICATION NUMBER: US 60/200,198
 ; PRIOR FILING DATE: 2000-04-28
 ; NUMBER OF SEQ ID NOS: 15
 ; SOFTWARE: Patentin version 3.1
 ; SEQ ID NO 6
 ; LENGTH: 419
 ; TYPE: PRT
 ; ORGANISM: Mus musculus
 ; US-09-843-905A-6

Query Match 95.0%; Score 2175.5; DB 3; Length 419;
 Best Local Similarity 95.0%; Pred. No. 3.1e-205;
 Matches 399; Conservative 10; Mismatches 10; Indels 1; Gaps 1;

QY 1 MFSPGQEHCAPNKEPVKYGELVVLGYNGALPNGDRGRKRSFALYKRPANGVKSTVH 60
 DB 1 MFSPGQEHCAPNKEPVKYGELVVLGYNGALPNGDRGRKRSFALYKRPANGVKSTVH 60
 QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVEYTHDKOTDMFOVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQASKAISCKGHSISYTLSSRNQTVVEYTHDKOTDMFOVGRSTESPIDFVVTDTI 120
 QY 121 SSGQNTDEAQTOSTISRFACRIYCDNREPYTARIIPAAGDSSKNITLGEKAKMKNPDG 180
 DB 121 SSGQNTDEAQTOSTISRFACRIYCDNREPYTARIIPAAGDSSKNITLGEKAKMKNPDG 180
 QY 181 HMDGLTNGVLVNHMPRGFTESQPGVWREISVCGDVYTLRETRSAQQRGLVSESTNVL 240
 DB 181 HMDGLTNGVLVNHMPRGFTESQPGVWREISVCGDVYTLRETRSAQQRGLVSESTNVL 240
 QY 241 QDQSLIDLCGATLLMRTADGLFHTPTQKHIALROBINARPOCPVGLNTLAPPSINRKE 300
 DB 241 QDQSLIDLCGATLLMRTADGLFHTPTQKHIALROBINARPOCPVGLNTLAPPSINRKE 300
 QY 301 VVEBKQPMAYLSCGHVGHYHNMGRSDTEANERECPCMCRTVGPVPLMLGCEAGFYVDAG 360
 DB 301 VVEBKQPMAYLSCGHVGHYHNMGRSDTEANERECPCMCRTVGPVPLMLGCEAGFYVDAG 360
 QY 361 PPTHAFTPCGHVSSEKSAKXWSQIPLPHGTHAFHAACPFCATQLVGBQNCIKLIFQGPID 420
 DB 361 PPTHAFTPCGHVSSEKSAKXWSQIPLPHGTHAFHAACPFCATQLVGBQNCIKLIFQGPID 420

RESULT 7

US-10-317-250-6
 ; Sequence 6, Application US/10317250
 ; Publication No. US20030165945A1
 ; GENERAL INFORMATION:
 ; APPLICANT: Bird, Timothy A.

APPLICANT: Cosman, David J.
APPLICANT: Li, Xiaoxia
TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
FILE REFERENCE: 2990-B
CURRENT APPLICATION NUMBER: US/10/317,250
CURRENT FILING DATE: 2002-12-11
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 419
TYPE: PRT
ORGANISM: Mus musculus
US-10-317-250-6

Query Match 95.0%; Score 2175.5; DB 4; Length 419;
Best Local Similarity 95.0%; Pred. No. 3.1e-205;
Matches 399; Conservative 10; Mismatches 10; Indels 1; Gaps 1;

QY 1 MFSPQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFPALYKRPYANGVKPSTVH 60
DB 1 MFSPQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFPALYKRPYANGVKPSTVH 60
QY 61 VISTPQASKAISCKGHSISYTLSSRQTVVVEYTHKDMDPVGSTSPIDFVVTDTI 120
DB 61 VISTPQASKAISCKGHSISYTLSSRQTVVVEYTHKDMDPVGSTSPIDFVVTDTI 120
QY 121 SSGQNTDEAQITOSTISRFACRIVCDBNEPYTARIFAAGFSSKNIFLGEKAKWKMPDG 180
DB 121 SSGQNTDEAQITOSTISRFACRIVCDBNEPYTARIFAAGFSSKNIFLGEKAKWKMPDG 180
QY 181 HMDGLTNGVLVNHPRGFTFESOPGWREISVCGGVYTLRETRSAOQRKLVESSTNVL 240
DB 181 HMDGLTNGVLVNHPRGFTFESOPGWREISVCGGVYTLRETRSAOQRKLVESSTNVL 240
QY 241 QDGLIDLCGATILMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
DB 241 QDGLIDLCGATILMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
QY 301 VVEKQPMAYLSCGHVGHYHNGHSDTEANERECPCRTVGPVPLMGEAGFYVDAG 360
DB 301 VVEKQPMAYLSCGHVGHYHNGHSDTEANERECPCRTVGPVPLMGEAGFYVDAG 360
QY 361 PPTHAFTPCGHVCSEKSAKYSQIPLPHGTHAFHAACPFCAQTQVGEONCIKILFQGPVD 420
DB 361 PPTHAFTPCGHVCSEKSAKYSQIPLPHGTHAFHAACPFCAQTQVGEONCIKILFQGPVD 420

RESULT 8
US-10-258-703-6
Sequence 6, Application US/10258703
Publication No. US20040034199A1
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
APPLICANT: Cosman, David J.
TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
FILE REFERENCE: 2990-US
CURRENT APPLICATION NUMBER: US/10/258,703
CURRENT FILING DATE: 2002-10-24
NUMBER OF SEQ ID NOS: 15
SOFTWARE: PatentIn version 3.1
SEQ ID NO 6
LENGTH: 419
TYPE: PRT
ORGANISM: Mus musculus
US-10-258-703-6

Query Match 95.0%; Score 2175.5; DB 4; Length 419;
Best Local Similarity 95.0%; Pred. No. 3.1e-205;
Matches 399; Conservative 10; Mismatches 10; Indels 1; Gaps 1;

QY 1 MFSPQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFPALYKRPYANGVKPSTVH 60
DB 1 MFSPQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFPALYKRPYANGVKPSTVH 60

QY 61 VISTPQASKAISCKGHSISYTLSSRQTVVVEYTHKDMDPVGSTSPIDFVVTDTI 120
DB 61 VISTPQASKAISCKGHSISYTLSSRQTVVVEYTHKDMDPVGSTSPIDFVVTDTI 120
QY 121 SSGQNTDEAQITOSTISRFACRIVCDBNEPYTARIFAAGFSSKNIFLGEKAKWKMPDG 180
DB 121 SSGQNTDEAQITOSTISRFACRIVCDBNEPYTARIFAAGFSSKNIFLGEKAKWKMPDG 180
QY 241 QDGLIDLCGATILMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
DB 241 QDGLIDLCGATILMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
QY 301 VVEKQPMAYLSCGHVGHYHNGHSDTEANERECPCRTVGPVPLMGEAGFYVDAG 360
DB 301 VVEKQPMAYLSCGHVGHYHNGHSDTEANERECPCRTVGPVPLMGEAGFYVDAG 360
QY 361 PPTHAFTPCGHVCSEKSAKYSQIPLPHGTHAFHAACPFCAQTQVGEONCIKILFQGPVD 420
DB 361 PPTHAFTPCGHVCSEKSAKYSQIPLPHGTHAFHAACPFCAQTQVGEONCIKILFQGPVD 420

RESULT 9
US-10-197-666A-134
Sequence 134, Application US/10197666A
Publication No. US20030092037A1
GENERAL INFORMATION:
APPLICANT: ASAH KASBI KASBI KASBI
TITLE OF INVENTION: Blk phosphorylation related gene
FILE REFERENCE: PH-1548US
CURRENT FILING DATE: 2002-11-18
CURRENT FILING DATE: 2001-11-18
PRIOR APPLICATION NUMBER: JP 2001-218204
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: JP 2001-263450
PRIOR FILING DATE: 2001-08-31
PRIOR APPLICATION NUMBER: JP 2002-012176
PRIOR FILING DATE: 2002-01-21
PRIOR APPLICATION NUMBER: US 60/305,884
PRIOR FILING DATE: 2001-07-18
PRIOR APPLICATION NUMBER: US 60/316,304
PRIOR FILING DATE: 2001-09-04
PRIOR APPLICATION NUMBER: US 60/350,027
PRIOR FILING DATE: 2002-01-23
NUMBER OF SEQ ID NOS: 156
SOFTWARE: PatentIn Ver. 2.1
SEQ ID NO 134
LENGTH: 419
TYPE: PRT
ORGANISM: Mus musculus
NAME/KEY: UNSURE
LOCATION: 357, 362
OTHER INFORMATION: unknown
US-10-197-666A-134

Query Match 92.2%; Score 2111.5; DB 4; Length 419;
Best Local Similarity 92.4%; Pred. No. 6.3e-199;
Matches 388; Conservative 12; Mismatches 19; Indels 1; Gaps 1;

QY 1 MFSPQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFPALYKRPYANGVKPSTVH 60
DB 1 MFSPQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRFPALYKRPYANGVKPSTVH 60
QY 61 VISTPQASKAISCKGHSISYTLSSRQTVVVEYTHKDMDPVGSTSPIDFVVTDTI 120
DB 61 VISTPQASKAISCKGHSISYTLSSRQTVVVEYTHKDMDPVGSTSPIDFVVTDTI 120
QY 121 SSGQNTDEAQITOSTISRFACRIVCDBNEPYTARIFAAGFSSKNIFLGEKAKWKMPDG 180

```

Db 121 SGGNED-AQITGSTISRFACRIICDNEPTAIIPAAGDSSKNITLGEAAAKMPDG 179
Qy 181 HMDGLTNGVLVNHPRGFTBESOPGWRIISVCGDYTTLTRSAOQKGLVESETNVL 240
Db 180 HMDGLTNGVLVNHPRGFTBESOPGWRIISVCGDYTTLTRSAOQKGLVESETNVL 239
Qy 241 QDGLSLIDLCATLMTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 300
Db 240 QDGLSLIDLCATLMTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 299
Qy 301 VVEKOPWATLSCGHVGHYHNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
Db 300 VVEKOPWATLSCGHVGHYHNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 359
Qy 361 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHFAHAACPFCATQLVGBONCIKLIPOGPID 420
Db 360 XXXHVFPCGHVCESEKSAKYSQIPLPHGTHFAHAACPFCATQLVGBONCIKLIPOGPID 419

```

```

RESULT 10
US-09-843-905A-4
; Sequence 4, Application US/09843905A
; Patent No. US20020168683A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Homo sapiens.
US-09-843-905A-4

```

```

Query Match 83.7%; Score 1917; DB 3; Length 418;
Best Local Similarity 81.4%; Pred. No. 9,4e-180;
Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;

```

```

Qy 1 MFSPGDEHCAPNKEPVYKGLVLYGNGALPNDGRGRKSRPALYKRPANGVKSTVH 60
Db 1 MFSPDQENH--PSKAPVYKGLVLYGNGSLPNDGRGRKSRPALYKRPANGVKSTVH 58
Qy 61 VISTPOASKAISCKGHSISYTLSSRNQTVVEYTHDKDTDMFOVGRSTESPIDFVVTDTI 120
Db 59 IACTPQAAKAIKNDQHSISYTLSSRAQTVVEYTHDSNTDMFOIGRSTESPIDFVVTDTV 118
Qy 121 SGGNDEAQTGSTISRFACRIICDNEPTAIIPAAGDSSKNITLGEAAAKMPDG 180
Db 119 PGSQNSDQSVOSTISRFACRIICERNPPTAIYAAGDSSKNITLGEAAAKMTSDG 178
Qy 181 HMDGLTNGVLVNHPRGFTBESOPGWRIISVCGDYTTLTRSAOQKGLVESETNVL 240
Db 179 QMDGLTNGVLVNHPRGFTBESOPGWRIISVCGDYTTLTRSAOQKGLVESETNVL 238
Qy 241 QDGLSLIDLCATLMTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 300
Db 239 QDGLSLIDLCATLMTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 298
Qy 301 VVEKOPWATLSCGHVGHYHNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
Db 299 VVEKOPWATLSCGHVGHYHNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 358
Qy 361 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHFAHAACPFCATQLVGBONCIKLIPOGPID 420
Db 359 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHFAHAACPFCATQLVGBONCIKLIPOGPID 418

```

```

RESULT 11
US-10-317-250-4
; Sequence 4, Application US/10317250
; Publication No. US20030165945A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; APPLICANT: Li, Xiaoxia
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-B
; CURRENT APPLICATION NUMBER: US/10/317,250
; CURRENT FILING DATE: 2002-12-11
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: Patentin version 3.1
; SEQ ID NO 4
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Homo sapiens
US-10-317-250-4

```

```

Query Match 83.7%; Score 1917; DB 4; Length 418;
Best Local Similarity 81.4%; Pred. No. 9,4e-180;
Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;

```

```

Qy 1 MFSPGDEHCAPNKEPVYKGLVLYGNGALPNDGRGRKSRPALYKRPANGVKSTVH 60
Db 1 MFSPDQENH--PSKAPVYKGLVLYGNGSLPNDGRGRKSRPALYKRPANGVKSTVH 58
Qy 61 VISTPOASKAISCKGHSISYTLSSRNQTVVEYTHDKDTDMFOVGRSTESPIDFVVTDTI 120
Db 59 IACTPQAAKAIKNDQHSISYTLSSRAQTVVEYTHDSNTDMFOIGRSTESPIDFVVTDTV 118
Qy 121 SGGNDEAQTGSTISRFACRIICDNEPTAIIPAAGDSSKNITLGEAAAKMPDG 180
Db 119 PGSQNSDQSVOSTISRFACRIICERNPPTAIYAAGDSSKNITLGEAAAKMTSDG 178
Qy 181 HMDGLTNGVLVNHPRGFTBESOPGWRIISVCGDYTTLTRSAOQKGLVESETNVL 240
Db 179 QMDGLTNGVLVNHPRGFTBESOPGWRIISVCGDYTTLTRSAOQKGLVESETNVL 238
Qy 241 QDGLSLIDLCATLMTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 300
Db 239 QDGLSLIDLCATLMTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 298
Qy 301 VVEKOPWATLSCGHVGHYHNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
Db 299 VVEKOPWATLSCGHVGHYHNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 358
Qy 361 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHFAHAACPFCATQLVGBONCIKLIPOGPID 420
Db 359 PPTHAFPCGHVCESEKSAKYSQIPLPHGTHFAHAACPFCATQLVGBONCIKLIPOGPID 418

```

```

RESULT 12
US-10-443-108-10
; Sequence 10, Application US/10443108
; Publication No. US20040005615A1
; GENERAL INFORMATION:
; APPLICANT: LI, JING
; APPLICANT: KU, DAVID
; APPLICANT: YANG, JIANXIN
; TITLE OF INVENTION: AMPLIFICATION AND OVEREXPRESSION OF ONCOGENES
; FILE REFERENCE: 38002-0049
; CURRENT APPLICATION NUMBER: US/10/443,108
; CURRENT FILING DATE: 2003-05-22
; PRIOR APPLICATION NUMBER: 60/398,099
; PRIOR FILING DATE: 2002-07-25
; PRIOR APPLICATION NUMBER: 60/382,606
; PRIOR FILING DATE: 2002-05-24
; NUMBER OF SEQ ID NOS: 90
; SOFTWARE: Patentin Ver. 2.1
; SEQ ID NO 10

```


LENGTH: 418
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-443-108-10

Query Match 83.7%; Score 1917; DB 4; Length 418;
 Best Local Similarity 81.4%; Pred. No. 9.4e-180;
 Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;

QY 1 MFSPGQEHCAPNKEPVYKGLVYLGYNGALPNGDRGRKSRPALYKRPKANGKPESTVH 60
 Db 1 MFSPDQENH--PSKAPVKYKGLVYLGYNGSLPNGDRGRKSRPALYKRPKANGKPESTVH 58
 QY 61 VISTPQASKAISCKGHSISYTLISRNQTVVVEYTHDKDTDMFQVGRSTESPIDFVYTDIT 120
 Db 59 IACTPQAKAISNDQHSISYTLISRAQTVVVEYTHDSNTDMFQIGRSTESPIDFVYTDIV 118
 QY 121 SSGSNTDEAQTOSTISRFACRIVCDBNEPYTARIIPAAGPDSKNITLGEKAKAKNPDG 180
 Db 119 PGSQNSDSTOSVOSTISRFACRIICERNPPTARIYAAGPDSKNITLGEKAKAKMTSDG 178
 QY 181 HMDGLTNGVLVNHPRGFTESQPGVWEISVCGDYVTLRETRSAQORGLVESETNVL 240
 Db 179 QMDGLTNGVLVNHPRNGFTEDSKPGIMREISVCGVFSLRERSAQORGLWEIETNQL 238
 QY 241 QDGLSLDLCGATLLMRTADGLFTPTQKHEALROBINARPOCPVGLNTLAFPSINRKE 300
 Db 239 QDGLSLDLCGATLLMRTABGLSHTPTVKHEALROBINARPOCPVGFNTLAFPSMKRKO 298
 QY 301 VVEKOPWAVLSCGHVGHYNNWGRSDTEANRECEPCRTVGPVYVPLMLGCEAGFYVDAG 360
 Db 299 VVEKOPWVVLNCGHVGHYNNWGNKBERDCKRECPMCRSVGVYVPLMLGCEAGFYVDAG 358
 QY 361 PPTHAFPCGHVCESEKAKYWSQIPLPHGTHFAACPCFCAQLVGEONCIKLIPOGPID 420
 Db 359 PPTHAFSPCGHVCESEKAKYWSQIPLPHGTHFAACPCFCAQLAGEGYIRLIFOGPLD 418

RESULT 13

US-10-258-703-4

Sequence 4, Application US/10258703
 Publication No. US20040034199A1

GENERAL INFORMATION:
 APPLICANT: Bird, Timothy A.
 APPLICANT: Coeman, David J.
 TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
 FILE REFERENCE: 2990-US
 CURRENT APPLICATION NUMBER: US/10/258, 703
 CURRENT FILING DATE: 2002-10-24
 NUMBER OF SEQ ID NOS: 15
 SOFTWARE: PatentIn version 3.1
 SEQ ID NO 4
 LENGTH: 418
 TYPE: PRT
 ORGANISM: Homo sapiens
 US-10-258-703-4

Query Match 83.7%; Score 1917; DB 4; Length 418;
 Best Local Similarity 81.4%; Pred. No. 9.4e-180;
 Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;

QY 1 MFSPGQEHCAPNKEPVYKGLVYLGYNGALPNGDRGRKSRPALYKRPKANGKPESTVH 60
 Db 1 MFSPDQENH--PSKAPVKYKGLVYLGYNGSLPNGDRGRKSRPALYKRPKANGKPESTVH 58
 QY 61 VISTPQASKAISCKGHSISYTLISRNQTVVVEYTHDKDTDMFQVGRSTESPIDFVYTDIT 120
 Db 59 IACTPQAKAISNDQHSISYTLISRAQTVVVEYTHDSNTDMFQIGRSTESPIDFVYTDIV 118
 QY 121 SSGSNTDEAQTOSTISRFACRIVCDBNEPYTARIIPAAGPDSKNITLGEKAKAKNPDG 180
 Db 119 PGSQNSDSTOSVOSTISRFACRIICERNPPTARIYAAGPDSKNITLGEKAKAKMTSDG 178

QY 181 HMDGLTNGVLVNHPRGFTESQPGVWEISVCGDYVTLRETRSAQORGLVESETNVL 240
 Db 179 QMDGLTNGVLVNHPRNGFTEDSKPGIMREISVCGVFSLRERSAQORGLWEIETNQL 238
 QY 241 QDGLSLDLCGATLLMRTADGLFTPTQKHEALROBINARPOCPVGLNTLAFPSINRKE 300
 Db 239 QDGLSLDLCGATLLMRTABGLSHTPTVKHEALROBINARPOCPVGFNTLAFPSMKRKO 298
 QY 301 VVEKOPWAVLSCGHVGHYNNWGRSDTEANRECEPCRTVGPVYVPLMLGCEAGFYVDAG 360
 Db 299 VVEKOPWVVLNCGHVGHYNNWGNKBERDCKRECPMCRSVGVYVPLMLGCEAGFYVDAG 358
 QY 361 PPTHAFPCGHVCESEKAKYWSQIPLPHGTHFAACPCFCAQLVGEONCIKLIPOGPID 420
 Db 359 PPTHAFSPCGHVCESEKAKYWSQIPLPHGTHFAACPCFCAQLAGEGYIRLIFOGPLD 418

RESULT 14

US-10-041-030-2

Sequence 2, Application US/10041030
 Publication No. US20020150934A1

GENERAL INFORMATION:
 APPLICANT: Powers, Scott
 APPLICANT: Mu, David
 APPLICANT: Xiang, Phil
 APPLICANT: Peng, Yue
 APPLICANT: Tularek Inc.
 TITLE OF INVENTION: Diagnosis and Treatment of Cancer Using Mammalian
 FILE REFERENCE: 018781-00681005
 CURRENT APPLICATION NUMBER: US/10/041, 030
 CURRENT FILING DATE: 2001-12-28
 PRIOR APPLICATION NUMBER: US 60/259, 502
 NUMBER OF SEQ ID NOS: 42
 SOFTWARE: PatentIn Ver. 2.1
 SEQ ID NO 2
 LENGTH: 418
 TYPE: PRT
 ORGANISM: Homo sapiens
 OTHER INFORMATION: human pellino 1
 US-10-041-030-2

Query Match 83.5%; Score 1913; DB 4; Length 418;
 Best Local Similarity 81.4%; Pred. No. 2.3e-179;
 Matches 342; Conservative 37; Mismatches 39; Indels 2; Gaps 1;

QY 1 MFSPGQEHCAPNKEPVYKGLVYLGYNGALPNGDRGRKSRPALYKRPKANGKPESTVH 60
 Db 1 MFSPDQENH--PSKAPVKYKGLVYLGYNGSLPNGDRGRKSRPALYKRPKANGKPESTVH 58
 QY 61 VISTPQASKAISCKGHSISYTLISRNQTVVVEYTHDKDTDMFQVGRSTESPIDFVYTDIT 120
 Db 59 IACTPQAKAISNDQHSISYTLISRAQTVVVEYTHDSNTDMFQIGRSTESPIDFVYTDIV 118
 QY 121 SSGSNTDEAQTOSTISRFACRIVCDBNEPYTARIIPAAGPDSKNITLGEKAKAKNPDG 180
 Db 119 PGSQNSDSTOSVOSTISRFACRIICERNPPTARIYAAGPDSKNITLGEKAKAKMTSDG 178
 QY 181 HMDGLTNGVLVNHPRGFTESQPGVWEISVCGDYVTLRETRSAQORGLVESETNVL 240
 Db 179 QMDGLTNGVLVNHPRNGFTEDSKPGIMREISVCGVFSLRERSAQORGLWEIETNQL 238
 QY 241 QDGLSLDLCGATLLMRTADGLFTPTQKHEALROBINARPOCPVGLNTLAFPSINRKE 300
 Db 239 QDGLSLDLCGATLLMRTABGLSHTPTVKHEALROBINARPOCPVGFNTLAFPSMKRKO 298
 QY 301 VVEKOPWAVLSCGHVGHYNNWGRSDTEANRECEPCRTVGPVYVPLMLGCEAGFYVDAG 360
 Db 299 VVEKOPWVVLNCGHVGHYNNWGNKBERDCKRECPMCRSVGVYVPLMLGCEAGFYVDAG 358
 QY 361 PPTHAFPCGHVCESEKAKYWSQIPLPHGTHFAACPCFCAQLVGEONCIKLIPOGPID 420

Db 359 PPTHAFFSCGHVSEKKTATWWSQILPHGHHTFHAAACPFCHQLAGEGYIRLLIFQSPLD 418

RESULT 15
US-09-843-905A-2
; Sequence 2, Application US/09843905A
; Patent No. US20020168683A1
; GENERAL INFORMATION:
; APPLICANT: Bird, Timothy A.
; APPLICANT: Cosman, David J.
; TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
; FILE REFERENCE: 2990-A
; CURRENT APPLICATION NUMBER: US/09/843,905A
; CURRENT FILING DATE: 2001-04-27
; PRIOR APPLICATION NUMBER: US 60/200,198
; PRIOR FILING DATE: 2000-04-28
; NUMBER OF SEQ ID NOS: 15
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 2
; LENGTH: 418
; TYPE: PRT
; ORGANISM: Mus musculus
US-09-843-905A-2

Query Match 83.5%; Score 1912; DB 3; Length 418;
Best Local Similarity 81.2%; Pred. No. 2.9e-179;
Matches 341; Conservative 39; Mismatches 38; Indels 2; Gaps 1;

QY 1 MFSPGQEHCAPNKEPVYKGLVVLGYNGALPNDGRGRKSRFALYKPRKANGVKPSTVH 60
DB 1 MFSPDQENH--PSKAPVKIGELIVLGYNGSLPNDGRGRKSRFALYKPRKANGVKSTVH 58
QY 61 VISTPQASKAISCKQHSISYTLSSRQTVVVEYTHDKDTMFOVGRSTESPIDFVVTDTI 120
DB 59 IACTPQAKAISNDQHSISYTLSSRAQTVVVEYTHDSNTDMFOIGRSTESPIDFVVTDTV 118
QY 121 SSGSNTDEAQTOSTISRFACRIYCDNRPYTARIFPAGFDSSKNITFLGEKAAKKNPDG 180
DB 119 PGSQNSDTSVOSTISRFACRIICERSPPPTARIYAAGFSSKNIFLGEKAAKMTSDG 178
QY 181 HMDGLTNGVLYMHPRGFTBESQGVWRREISVCGDYVTLRETPSAOQRKXIVSEETNVL 240
DB 179 QMDGLTNGVLYMHPRGFTBESQGVWRREISVCGDYVTLRETPSAOQRKXIVSEETNVL 238
QY 241 QDSGLIDLCGATLIMRTADGLFHTPTOKHLEALROEINARPOCPVGIANLTFPSINRKE 300
DB 239 QDSGLIDLCGATLIMRTABGLSHPTVYKHEALROEINARPOCPVGIANLTFPSMKRKD 298
QY 301 VVEBKQPMAYLSCGHVGHYTNMGHRSPTENARECEPMCRVTGYPVPLMLGCEAGFYVDAG 360
DB 299 VVEBKQPMVYVINGCHVGHYTNMGHRSPTENARECEPMCRVTGYPVPLMLGCEAGFYVDAG 358
QY 361 PPTHAFFSCGHVSEKSAKYWSQILPHGHHTFHAAACPFCHQLAGEGYIRLLIFQSPID 420
DB 359 PPTHAFFSCGHVSEKKTATWWSQILPHGHHTFHAAACPFCHQLAGEGYIRLLIFQSPID 418

Search completed: March 1, 2006, 21:04:32
Job time: 168 secs

Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2006, 21:00:16 ; Search time 47 Seconds

(without alignments)
738.804 Million cell updates/sec

Title: US-10-041-030-4

Perfect score: 2290

Sequence: 1 MFSPGQEHKAPKPEVKYK.....ATLVGEQNCIKILPGPID 420

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 572060 seqs, 82675679 residues

Total number of hits satisfying chosen parameters: 572060

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Issued Patents AA:*
1: /cgn2_6/ptodata/1/iaa/5 COMB.pep:*
2: /cgn2_6/ptodata/1/iaa/6 COMB.pep:*
3: /cgn2_6/ptodata/1/iaa/H COMB.pep:*
4: /cgn2_6/ptodata/1/iaa/PCUS COMB.pep:*
5: /cgn2_6/ptodata/1/iaa/RE COMB.pep:*
6: /cgn2_6/ptodata/1/iaa/backfile1.pep:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | DB ID | Description |
|------------|--------|-------------|--------|----------------------|--------------------|
| 1 | 2283 | 99.7 | 420 | US-09-843-905A-8 | Sequence 8, Appl |
| 2 | 2175.5 | 95.0 | 419 | US-09-843-905A-6 | Sequence 6, Appl |
| 3 | 1917 | 83.7 | 418 | US-09-843-905A-4 | Sequence 4, Appl |
| 4 | 1912 | 83.5 | 418 | US-09-843-905A-2 | Sequence 2, Appl |
| 5 | 1642 | 71.7 | 445 | US-09-843-905A-12 | Sequence 12, Appl |
| 6 | 1482.5 | 64.7 | 406 | US-10-104-047-3928 | Sequence 3928, Ap |
| 7 | 1302 | 56.9 | 424 | US-09-843-905A-13 | Sequence 13, Appl |
| 8 | 1235 | 53.9 | 455 | US-09-843-905A-14 | Sequence 14, Appl |
| 9 | 903 | 39.4 | 458 | US-09-843-905A-15 | Sequence 15, Appl |
| 10 | 104 | 4.5 | 710 | US-09-252-991A-25700 | Sequence 25700, A |
| 11 | 100 | 4.4 | 2042 | US-09-077-098A-6 | Sequence 6, Appl |
| 12 | 100 | 4.4 | 2042 | US-10-192-584-6 | Sequence 6, Appl |
| 13 | 96 | 4.2 | 717 | US-09-949-016-6500 | Sequence 6500, Ap |
| 14 | 95.5 | 4.2 | 978 | US-10-198-452A-485 | Sequence 485, App |
| 15 | 95.5 | 4.2 | 978 | US-10-197-220-65 | Sequence 456, Appl |
| 16 | 95.5 | 4.2 | 979 | US-09-438-185A-456 | Sequence 456, App |
| 17 | 94 | 4.1 | 2324 | US-08-283-857-1 | Sequence 1, Appl |
| 18 | 94 | 4.1 | 2324 | PCT-US95-09819-1 | Sequence 1, Appl |
| 19 | 93.5 | 4.1 | 727 | US-09-489-039A-14168 | Sequence 14168, A |
| 20 | 93.5 | 4.1 | 803 | US-09-543-681A-5399 | Sequence 5399, Ap |
| 21 | 93.5 | 4.1 | 1702 | US-09-854-133-834 | Sequence 434, App |
| 22 | 92 | 4.0 | 775 | US-09-949-016-8030 | Sequence 8030, Ap |
| 23 | 92 | 4.0 | 835 | US-09-949-016-7974 | Sequence 7974, Ap |
| 24 | 90 | 3.9 | 2231 | US-08-153-799-16 | Sequence 16, Appl |
| 25 | 88.5 | 3.9 | 558 | US-10-037-311B-1 | Sequence 1, Appl |
| 26 | 88.5 | 3.9 | 1104 | US-08-337-832-5 | Sequence 5, Appl |
| 27 | 88.5 | 3.9 | 1104 | US-08-828-584-5 | Sequence 5, Appl |

ALIGNMENTS

| | | | | | | |
|----|------|-----|------|---|----------------------|--------------------|
| 28 | 87.5 | 3.8 | 426 | 2 | US-09-252-991A-32215 | Sequence 32215, A |
| 29 | 87 | 3.8 | 453 | 2 | US-09-949-016-10252 | Sequence 10252, A |
| 30 | 87 | 3.8 | 254 | 1 | US-08-576-626A-32 | Sequence 32, Appl |
| 31 | 86.5 | 3.8 | 266 | 2 | US-10-104-047-3891 | Sequence 3891, Ap |
| 32 | 86.5 | 3.8 | 498 | 2 | US-09-248-796A-19085 | Sequence 19085, A |
| 33 | 86.5 | 3.8 | 1371 | 2 | US-09-302-540-16024 | Sequence 16024, A |
| 34 | 86 | 3.8 | 1454 | 2 | US-09-949-016-7599 | Sequence 7599, Ap |
| 35 | 86 | 3.8 | 1454 | 2 | US-09-949-016-7600 | Sequence 7600, Ap |
| 36 | 86 | 3.8 | 2327 | 6 | 5455158-1 | Patent No. 5455158 |
| 37 | 86 | 3.8 | 2355 | 2 | US-10-360-101-235 | Sequence 235, App |
| 38 | 86 | 3.8 | 2386 | 1 | US-09-016-366A-12 | Sequence 12, Appl |
| 39 | 86 | 3.8 | 2386 | 2 | US-09-961-403-1 | Sequence 1, Appl |
| 40 | 86 | 3.8 | 2446 | 1 | US-08-551-356-2 | Sequence 2, Appl |
| 41 | 86 | 3.8 | 2446 | 4 | PCT-US93-12687-2 | Sequence 2, Appl |
| 42 | 85.5 | 3.7 | 525 | 1 | US-08-160-861-4 | Sequence 4, Appl |
| 43 | 85.5 | 3.7 | 572 | 1 | US-08-160-861-3 | Sequence 3, Appl |
| 44 | 85.5 | 3.7 | 915 | 2 | US-09-817-514A-6 | Sequence 6, Appl |
| 45 | 85.5 | 3.7 | 1073 | 2 | US-09-949-016-9771 | Sequence 9771, Ap |

| | | | | | | |
|--|-----|--|-----|--|-----|--|
| RESULT 1 | | | | | | |
| US-09-843-905A-8 | | | | | | |
| Sequence 8, Application US/09843905A | | | | | | |
| Patent No. 6703487 | | | | | | |
| GENERAL INFORMATION: | | | | | | |
| APPLICANT: Bird, Timothy A. | | | | | | |
| TITLE OR INVENTION: HUMAN PELLINO POLYPEPTIDES | | | | | | |
| FILE REFERENCE: 2999A | | | | | | |
| CURRENT APPLICATION NUMBER: US/09/843,905A | | | | | | |
| CURRENT FILING DATE: 2001-04-27 | | | | | | |
| PRIOR APPLICATION NUMBER: US-60/200,198 | | | | | | |
| PRIOR FILING DATE: 2000-04-28 | | | | | | |
| NUMBER OF SEQ ID NOS: 15 | | | | | | |
| SOFTWARE: Patent version 3.1 | | | | | | |
| SEQ ID NO 8 | | | | | | |
| LENGTH: 420 | | | | | | |
| TYPE: PRT | | | | | | |
| ORGANISM: Homo sapiens | | | | | | |
| US-09-843-905A-8 | | | | | | |
| Query Match | | | | | | |
| Best Local Similarity 99.8%; Pred. No. 8.3e-238; | | | | | | |
| Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0; | | | | | | |
| QY | 1 | MFSPGQEHKAPKPEVKYK | 1 | MFSPGQEHKAPKPEVKYK | 1 | MFSPGQEHKAPKPEVKYK |
| DB | 1 | MFSPGQEHKAPKPEVKYK | 1 | MFSPGQEHKAPKPEVKYK | 1 | MFSPGQEHKAPKPEVKYK |
| QY | 61 | VISTPQASKISKGGHISYTLRNQTVVETHTKQDMDPVGSTESPIPVVTDI | 120 | VISTPQASKISKGGHISYTLRNQTVVETHTKQDMDPVGSTESPIPVVTDI | 120 | VISTPQASKISKGGHISYTLRNQTVVETHTKQDMDPVGSTESPIPVVTDI |
| DB | 61 | VISTPQASKISKGGHISYTLRNQTVVETHTKQDMDPVGSTESPIPVVTDI | 120 | VISTPQASKISKGGHISYTLRNQTVVETHTKQDMDPVGSTESPIPVVTDI | 120 | VISTPQASKISKGGHISYTLRNQTVVETHTKQDMDPVGSTESPIPVVTDI |
| QY | 121 | SGSQNTDEAQTOSTISRFACRIVCDRNEBYTARIFAGDSKNIPLFGKAKAKNPDG | 180 | SGSQNTDEAQTOSTISRFACRIVCDRNEBYTARIFAGDSKNIPLFGKAKAKNPDG | 180 | SGSQNTDEAQTOSTISRFACRIVCDRNEBYTARIFAGDSKNIPLFGKAKAKNPDG |
| DB | 121 | SGSQNTDEAQTOSTISRFACRIVCDRNEBYTARIFAGDSKNIPLFGKAKAKNPDG | 180 | SGSQNTDEAQTOSTISRFACRIVCDRNEBYTARIFAGDSKNIPLFGKAKAKNPDG | 180 | SGSQNTDEAQTOSTISRFACRIVCDRNEBYTARIFAGDSKNIPLFGKAKAKNPDG |
| QY | 181 | HMGLTTNGVTVNHPFGFTBESQPGVWRISVCGVYTLRFRSKAQKGLVSESTNVL | 240 | HMGLTTNGVTVNHPFGFTBESQPGVWRISVCGVYTLRFRSKAQKGLVSESTNVL | 240 | HMGLTTNGVTVNHPFGFTBESQPGVWRISVCGVYTLRFRSKAQKGLVSESTNVL |
| DB | 181 | HMGLTTNGVTVNHPFGFTBESQPGVWRISVCGVYTLRFRSKAQKGLVSESTNVL | 240 | HMGLTTNGVTVNHPFGFTBESQPGVWRISVCGVYTLRFRSKAQKGLVSESTNVL | 240 | HMGLTTNGVTVNHPFGFTBESQPGVWRISVCGVYTLRFRSKAQKGLVSESTNVL |
| QY | 241 | QDSGLIDLCATLTMTADGLFTPTQKHBAVROEINARPCQVGLMTLAPPSINRKE | 300 | QDSGLIDLCATLTMTADGLFTPTQKHBAVROEINARPCQVGLMTLAPPSINRKE | 300 | QDSGLIDLCATLTMTADGLFTPTQKHBAVROEINARPCQVGLMTLAPPSINRKE |
| DB | 241 | QDSGLIDLCATLTMTADGLFTPTQKHBAVROEINARPCQVGLMTLAPPSINRKE | 300 | QDSGLIDLCATLTMTADGLFTPTQKHBAVROEINARPCQVGLMTLAPPSINRKE | 300 | QDSGLIDLCATLTMTADGLFTPTQKHBAVROEINARPCQVGLMTLAPPSINRKE |
| QY | 301 | VVEKQPMWATLSCGHVGHYHNGHRSDDTEANRECECMCTVGGVYVLMGCEAGFYVDAG | 360 | VVEKQPMWATLSCGHVGHYHNGHRSDDTEANRECECMCTVGGVYVLMGCEAGFYVDAG | 360 | VVEKQPMWATLSCGHVGHYHNGHRSDDTEANRECECMCTVGGVYVLMGCEAGFYVDAG |
| DB | 301 | VVEKQPMWATLSCGHVGHYHNGHRSDDTEANRECECMCTVGGVYVLMGCEAGFYVDAG | 360 | VVEKQPMWATLSCGHVGHYHNGHRSDDTEANRECECMCTVGGVYVLMGCEAGFYVDAG | 360 | VVEKQPMWATLSCGHVGHYHNGHRSDDTEANRECECMCTVGGVYVLMGCEAGFYVDAG |
| QY | 361 | PPHTAFTPCGHVCESEKSAKTSQIPLPHGTHAFAACFPATQVGEQNCIKILPGPID | 420 | PPHTAFTPCGHVCESEKSAKTSQIPLPHGTHAFAACFPATQVGEQNCIKILPGPID | 420 | PPHTAFTPCGHVCESEKSAKTSQIPLPHGTHAFAACFPATQVGEQNCIKILPGPID |
| DB | 361 | PPHTAFTPCGHVCESEKSAKTSQIPLPHGTHAFAACFPATQVGEQNCIKILPGPID | 420 | PPHTAFTPCGHVCESEKSAKTSQIPLPHGTHAFAACFPATQVGEQNCIKILPGPID | 420 | PPHTAFTPCGHVCESEKSAKTSQIPLPHGTHAFAACFPATQVGEQNCIKILPGPID |

Db 361 PPTHAFTPCGHVSEKSAKATWSQIPLPHGTHAFHAACPFCAATQLVGEONCIKILFQGPID 420

RESULT 2

US-09-843-905A-6
Sequence 6, Application US/09843905A

Patent No. 6703487
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
APPLICANT: Cosman, David J.
TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
FILE REFERENCE: 2990-A
CURRENT APPLICATION NUMBER: US/09/843,905A
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/200,198
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin version 3.1
SEQ ID NO 6
LENGTH: 419
TYPE: PRT
ORGANISM: Mus musculus
US-09-843-905A-6

Query Match 95.0%; Score 2175.5; DB 2; Length 419;
Best Local Similarity 95.0%; Pred. No. 3.5e-226;
Matches 399; Conservative 10; Mismatches 10; Indels 1; Gaps 1;

QY 1 MFSPGQEBHCAPNKEPVYKYGELVVLGYNGALPNDGRGRKSRPALYKRPANGVKSTVH 60
Db 1 MFSPGQEBHCAPNKEPVYKYGELVVLGYNGALPNDGRGRKSRPALYKRYTASGVKPESTIH 60
QY 61 VISTPOASKAISCKGHSISYTLNRQTVVEYTHKODTDMFOVGRSTESPIDFVVTDTI 120
Db 61 MVTSPQASKAISCKGHSISYTLNRQTVVEYTHKODTDMFOVGRSTESPIDFVVTDTV 120
QY 121 SGGQNTDEAQITQSTISRPAACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKMKNPDG 180
Db 121 SGGQNTDEAQITQSTISRPAACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKMKNPDG 179
QY 181 HMDGLTTNGVLVNHPRGFTBESQPGVWREISVCGDYVTLRETRSAQGRKLVESSTNVL 240
Db 181 HMDGLTTNGVLVNHPRGFTBESQPGVWREISVCGDYVTLRETRSAQGRKLVESSTNVL 239
QY 241 QDQSLIDLCGATLLMRTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 300
Db 241 QDQSLIDLCGATLLMRTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 299
QY 301 VVEKOPMAYLSCGHVGHYNNWGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
Db 301 VVEKOPMAYLSCGHVGHYNNWGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 359
QY 361 PPTHAFTPCGHVSEKSAKATWSQIPLPHGTHAFHAACPFCAATQLVGEONCIKILFQGPID 420
Db 361 PPTHAFTPCGHVSEKSAKATWSQIPLPHGTHAFHAACPFCAATQLVGEONCIKILFQGPID 419

RESULT 3

US-09-843-905A-4
Sequence 4, Application US/09843905A

Patent No. 6703487
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
APPLICANT: Cosman, David J.
TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
FILE REFERENCE: 2990-A
CURRENT APPLICATION NUMBER: US/09/843,905A
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/200,198
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin version 3.1

SEQ ID NO 4
LENGTH: 418
TYPE: PRT
ORGANISM: Homo sapiens
US-09-843-905A-4

Query Match 83.7%; Score 1917; DB 2; Length 418;
Best Local Similarity 81.4%; Pred. No. 3.2e-198;
Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;

QY 1 MFSPGQEBHCAPNKEPVYKYGELVVLGYNGALPNDGRGRKSRPALYKRPANGVKSTVH 60
Db 1 MFSPGQEBHCAPNKEPVYKYGELVVLGYNGALPNDGRGRKSRPALYKRPANGVKSTVH 58
QY 61 VISTPOASKAISCKGHSISYTLNRQTVVEYTHKODTDMFOVGRSTESPIDFVVTDTI 120
Db 59 IACTPOAAKAIKNDQHSISYTLNRQTVVEYTHKODTDMFOVGRSTESPIDFVVTDTV 118
QY 121 SGGQNTDEAQITQSTISRPAACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKMKNPDG 180
Db 119 PGQNSNDTQSVQSTISRPAACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKMKNPDG 178
QY 181 HMDGLTTNGVLVNHPRGFTBESQPGVWREISVCGDYVTLRETRSAQGRKLVESSTNVL 240
Db 179 QMDGLTTNGVLVNHPRGFTBESQPGVWREISVCGDYVTLRETRSAQGRKLVESSTNVL 238
QY 241 QDQSLIDLCGATLLMRTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 300
Db 239 QDQSLIDLCGATLLMRTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 298
QY 301 VVEKOPMAYLSCGHVGHYNNWGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
Db 299 VVEKOPMAYLSCGHVGHYNNWGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 358
QY 361 PPTHAFTPCGHVSEKSAKATWSQIPLPHGTHAFHAACPFCAATQLVGEONCIKILFQGPID 420
Db 359 PPTHAFTPCGHVSEKSAKATWSQIPLPHGTHAFHAACPFCAATQLVGEONCIKILFQGPID 418

RESULT 4

US-09-843-905A-2
Sequence 2, Application US/09843905A

Patent No. 6703487
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
APPLICANT: Cosman, David J.
TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
FILE REFERENCE: 2990-A
CURRENT APPLICATION NUMBER: US/09/843,905A
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/200,198
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin version 3.1
SEQ ID NO 2
LENGTH: 418
TYPE: PRT
ORGANISM: Mus musculus
US-09-843-905A-2

Query Match 83.5%; Score 1912; DB 2; Length 418;
Best Local Similarity 81.2%; Pred. No. 1.1e-197;
Matches 341; Conservative 39; Mismatches 38; Indels 2; Gaps 1;

QY 1 MFSPGQEBHCAPNKEPVYKYGELVVLGYNGALPNDGRGRKSRPALYKRPANGVKSTVH 60
Db 1 MFSPGQEBHCAPNKEPVYKYGELVVLGYNGALPNDGRGRKSRPALYKRPANGVKSTVH 58
QY 61 VISTPOASKAISCKGHSISYTLNRQTVVEYTHKODTDMFOVGRSTESPIDFVVTDTI 120
Db 59 IACTPOAAKAIKNDQHSISYTLNRQTVVEYTHKODTDMFOVGRSTESPIDFVVTDTV 118
QY 121 SGGQNTDEAQITQSTISRPAACRIVCDRNEPYTARIFAAGFDSKNIFLGEKAAKMKNPDG 180

```

119 PGSGNSNDTQSVGSTITRPFACRIRICERSPPPTAITYAAGPSSKNITLGEAAAKMTSDG 178
181 HMDGLTNGVLVWHPRGFTBESOPGVWREISVCGDYVTLRETRSAOQKGLVESETNVL 240
179 QMGDLTNGVLVWHPRGFTBESOPGVWREISVCGDYVTLRETRSAOQKGLVESETNVL 238
241 QDPSLIDLCATLMTWRTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 300
239 QDPSLIDLCATLMTWRTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 298
301 VVEKOPWATYISCGVHYHNMGRSDTEANERECPCMTGVPYVPLMLGCEAGFYVDAG 360
299 VDEKOPWATYISCGVHYHNMGRSDTEANERECPCMTGVPYVPLMLGCEAGFYVDAG 358
361 PPTHAFTPCGHVCSKSAKTYWSQIPLPHGTHAFHAACPCFATOLVGEONCIKLIPOGPID 420
359 PPTHAFTPCGHVCSKSAKTYWSQIPLPHGTHAFHAACPCFATOLVGEONCIKLIPOGPID 418

```

RESULT 5
US-09-843-905A-12
Sequence 12, Application US/09843905A
Patent No. 6703487
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
APPLICANT: Cosman, David J.
TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
FILE REFERENCE: 2990-A
CURRENT APPLICATION NUMBER: US/09/843,905A
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/200,198
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin version 3.1
SEQ ID NO 12
LENGTH: 445
TYPE: PRT
ORGANISM: Homo sapiens
US-09-843-905A-12

Query Match 71.7%; Score 1642; DB 2; Length 445;
Best Local Similarity 71.4%; Pred. No. 2e-168;
Matches 299; Conservative 42; Mismatches 74; Indels 4; Gaps 3;

```

3 SPQGEHCANKEPVYKGLVLTGYNALPNDGRGRKSRPALKYKPKANGVKSPTVHY 62
30 SPG--EDAQGEPIKTKGELIVLGYNGCLASGDKRRSRALSRSHANGVKKDVVHHI 87
63 STPOASFAISCKGSHISYTLRSNQTUVVETHTDKDTMPQVGRSTESPIDFVVTDTISG 122
88 STPLVSKALSNRQSHISYTLRSNQTUVVETHTDKDTMPQVGRSTESPIDFVVTDTISG 147
123 SONTDEAQTOSTISRFACRIVCDRNEPYTARIFAAGPDSKNIPLGEKAKKKNPDGM 182
148 G-GAEGSPASQSTISRYACRLCDRPPYARIYAAGPDSNSNIFLGERAAKMTPTDGLM 206
183 DGLTTNGVLVWHPRGFTBESOPGVWREISVCGDYVTLRETRSAOQKGLVESETNVL 242
207 DGLTTNGVLVWHPRGFTBESOPGVWREISVCGDYVTLRETRSAOQKGLVESETNVL 266
243 GSLIDLCATLMTWRTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 302
267 GSLIDLCATLMTWRTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 326
303 BEK--QPAVYISCGVHYHNMGRSDTEANERECPCMTGVPYVPLMLGCEAGFYVDAG 361
327 PDKQOPWATYISCGVHYHNMGRSDTEANERECPCMTGVPYVPLMLGCEAGFYVDAG 386
362 PTHAFTPCGHVCSKSAKTYWSQIPLPHGTHAFHAACPCFATOLVGEONCIKLIPOGPID 420
387 PTHAFTPCGHVCSKSAKTYWSQIPLPHGTHAFHAACPCFATOLVGEONCIKLIPOGPID 445

```

```

US-10-104-047-3928
Sequence 3928, Application US/10104047
Patent No. 6943241
GENERAL INFORMATION:
APPLICANT: HELIX RESEARCH INSTITUTE
TITLE OF INVENTION: No. 6943241el full length cdna
FILE REFERENCE: HI-A0105
CURRENT APPLICATION NUMBER: US/10/104,047
CURRENT FILING DATE: 2002-03-25
PRIOR APPLICATION NUMBER:
PRIOR FILING DATE:
NUMBER OF SEQ ID NOS: 4096
SOFTWARE: Patentin Ver. 2.1
SEQ ID NO 3928
LENGTH: 406
TYPE: PRT
ORGANISM: Homo sapiens
US-10-104-047-3928

```

Query Match 64.7%; Score 1482.5; DB 2; Length 406;
Best Local Similarity 65.6%; Pred. No. 3e-151;
Matches 275; Conservative 36; Mismatches 65; Indels 43; Gaps 4;

```

3 SPQGEHCANKEPVYKGLVLTGYNALPNDGRGRKSRPALKYKPKANGVKSPTVHY 62
30 SPG--EDAQGEPIKTKGELIVLGYNGCLASGDKRRSRALSRSHANGVKKDVVHHI 87
63 STPOASFAISCKGSHISYTLRSNQTUVVETHTDKDTMPQVGRSTESPIDFVVTDTISG 122
49 STPLVSKALSNRQSHISYTLRSNQTUVVETHTDKDTMPQVGRSTESPIDFVVTDTISG 108
123 SONTDEAQTOSTISRFACRIVCDRNEPYTARIFAAGPDSKNIPLGEKAKKKNPDGM 182
109 G-GAEGSPASQSTISRYACRLCDRPPYARIYAAGPDSNSNIFLGERAAKMTPTDGLM 167
183 DGLTTNGVLVWHPRGFTBESOPGVWREISVCGDYVTLRETRSAOQKGLVESETNVL 242
168 DGLTTNGVLVWHPRGFTBESOPGVWREISVCGDYVTLRETRSAOQKGLVESETNVL 227
243 GSLIDLCATLMTWRTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 302
228 GSLIDLCATLMTWRTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 287
303 BEK--QPAVYISCGVHYHNMGRSDTEANERECPCMTGVPYVPLMLGCEAGFYVDAG 361
288 PDKQOPWATYISCGVHYHNMGRSDTEANERECPCMTGVPYVPLMLGCEAGFYVDAG 347
362 PTHAFTPCGHVCSKSAKTYWSQIPLPHGTHAFHAACPCFATOLVGEONCIKLIPOGPID 420
348 PTHAFTPCGHVCSKSAKTYWSQIPLPHGTHAFHAACPCFATOLVGEONCIKLIPOGPID 406

```

RESULT 7
US-09-843-905A-13
Sequence 13, Application US/09843905A
Patent No. 6703487
GENERAL INFORMATION:
APPLICANT: Bird, Timothy A.
APPLICANT: Cosman, David J.
TITLE OF INVENTION: HUMAN PELLINO POLYPEPTIDES
FILE REFERENCE: 2990-A
CURRENT APPLICATION NUMBER: US/09/843,905A
CURRENT FILING DATE: 2001-04-27
PRIOR APPLICATION NUMBER: US 60/200,198
PRIOR FILING DATE: 2000-04-28
NUMBER OF SEQ ID NOS: 15
SOFTWARE: Patentin version 3.1
SEQ ID NO 13
LENGTH: 424
TYPE: PRT
ORGANISM: Drosophila melanogaster

US-09-843-905A-13

| | | | | |
|-----------------------|------------------|---------------------|------------|-------------|
| Query Match | 56.9%; | Score 1302; | DB 2; | Length 424; |
| Best Local Similarity | 59.6%; | Pred. No. 1.1e-131; | | |
| Matches 249; | Conservative 65; | Mismatches 92; | Indels 12; | Gaps 7; |

| | | | |
|----|-----|---|-----|
| Qy | 5 | GOEHQAPNKEPXYKGEVLVLYNGALPLMGDGRKSKSPALYKRPKANGVSPVWHYST | 64 |
| Db | 17 | GGDGH---DKPRRLRYELVILGNGYLPQDGRRRSKFYVLLKKTMSGVKRSKHTYVQS | 73 |
| Qy | 65 | POASKASCKGHSISYTLRSNQTVVEXYHDKDMDFOVGRSTESPIDFVYDTTISGSQ | 124 |
| Db | 74 | PQSKATILDANGHSISYTLRSNQAVIYVEKEDTETDMFOVGRSSSPIDFVYDMLTGDK | 133 |
| Qy | 125 | NMDEADITGSTISRFACRIVCDRNEPYTKATIPAGFDSSKNTIFGEKRAKKAPDGMHG | 184 |
| Db | 134 | K--DAKVMSITSRFCRILIVRCEPAKARIAPAGFDSRNIPFGEKATKQD--NVEIDG | 190 |
| Qy | 185 | LTTNGVLYNMHPGCGF--EESQGVWRREISVCGDVYTLLETSSAQRGLVSENNYLODG | 243 |
| Db | 191 | LTTNGVLYNMHPKCSFGCGNAGKGLMRECSVGGDVYTLLETSSAQQKGPYIDECNIILODG | 250 |
| Qy | 244 | SLIDLGCATLLWRITDGLFHTPTOKIETAROEIJAARPOCPVGLNTLAPF--SINRKEVY | 302 |
| Db | 251 | TLIDLCGATLLWRMSAGLQHSPTKDDLETLIDAIINAGRPQCPVGLNTLVIPRKNIGDQV | 310 |
| Qy | 303 | EEKQPMAYIISCGHVGSTYHNGHRSPTDEANERECPCRTGVGPVPLMTGCEAGFYVDAGPP | 362 |
| Db | 311 | --NQPVYIYINCGVHQHDMGODENTGA--RRCPCIELGEPVTLTLCMGLBAPFYVDGAP | 366 |
| Qy | 363 | THAFPCGIVCSKSKAKVWSQIPRLHGTHAFLAARPCATQVLVEQNCIKILIPGPIID | 420 |
| Db | 367 | TYAFNCGHNAIEKTYKVYANVEIIPHGNGFAPVCPATPIPDGATGYIKILIPDNDID | 424 |

RESULT 8
 US-09-843-905A-14
 : Sequence 14, Application US/09843905A
 : Patent No. 6703487
 : GENERAL INFORMATION:
 :
 : APPLICANT: Bldd, Timothy A.
 : APPLICANT: Cosman, David U.
 : TITLE OF INVENTION: HUMAN PELLINO POLYPEPT
 : FILE REFERENCE: 2990-A
 :
 : CURRENT APPLICATION NUMBER: US/09/843,905A
 :
 : CURRENT FILING DATE: 2001-04-27
 : PRIOR APPLICATION NUMBER: US 60/200,198
 : PRIOR FILING DATE: 2000-04-28
 : NUMBER OF SEQ ID NOS: 15
 : SOFTWARE: PatentIn version 3.1
 : SEQ ID NO 14
 :
 : LENGTH: 455
 : TYPE: PRT
 : ORGANISM: Cloma intestinalis
 : US-09-843-905A-14

| Query Match | 53.9% | Score 1235 | DB 2 | Length 455 |
|-------------|-------|------------|------|------------|
|-------------|-------|------------|------|------------|

QY QEEHCAPKEPVKAGELVVLVYNGALPAGDGRGRSRFALYKRPANVQKSTYVH-VYST 64
 Db QEDAPKEBEGDIYIGQLVLVLNGOLPFGDGKRRSCTFLRKRVATGVEBDDQVYQK 96
 QY POASKAISCKOHSSISYTLSSRNOQVWEVETDKIDDMFQVGRSTESPIDPVYDTISGS- 123
 Db ASHSETFLSKOHHSVSTYLPRS-VVVVYVHADNDSDMQIGRSTEEPIDPVLMLDEAGS 155
 QY 124 ---ONTDEAQITOSTISRFACRIVCDRNEPYTARIFAAGFSSKNIPIGEKAKA---KN 177
 Db 156 IPTNKKPQTKQKSTISRFACRIVCDRNEPYTARIFAAGFSSKNIPIGEKAAPKTTTQN 215
 QY 178 PDGMDGLTTNGVLVMAHPRGGFTBESQGVWEISVCGDVYTLRSTRQAQKGLVESET 237

| | | | |
|----|-----|--|-----|
| Db | 216 | GKKIIIGLTITNGVLINQPKNGFSESTPTPOKETSVCQNIITQLAEKSSAQQLPGRIMEDN | 275 |
| Oy | 238 | NVYQDSLLIDCGATITLMTADGLFHTPTQKITEAKQETIAAARPCQPVGLNTIAFSSIN | 297 |
| Db | 276 | NTLVNGLTLDLCGATILMRSSSHERCMPPLHIDELHKLVLRPQCQVGLTTIAFPRRS | 335 |
| Oy | 298 | RKEVVEKQPMAYLSCGHVGHYTHNGHRSDTENAREECPCMCRTGPPVPLMLGCEAGFYV | 357 |
| Db | 336 | KAKETEEKPMWYLQCGHVGRIENGQGE--BEIRICPLCRSVGKVVPLVWGGEPAFYV | 392 |
| Oy | 358 | DAGPPTHAQPCGHVSESEKAYKMSQITPLPHGTNLFHAACGFCAYTOLVGESENCIKLIFQG | 417 |
| Db | 393 | DIERPEYCFVPCGHVCSQKTAIYVSGTALPHGTQAVSACFCFCAPILEGDGVKYLIFQG | 452 |
| Oy | 418 | PID 420 | |
| Db | 453 | PLD 455 | |

```

RESULT 9
US-09-843-905A-15
: Sequence 15, Application US/09843905A
: Patent No. 6703487
: GENERAL INFORMATION:
: APPLICANT: Bird, Timothy A.
: APPLICANT: Cosman, David U.
: TITLE OF INVENTION: HUMAN PELLINO POLYPEPT
: FILE REFERENCE: 2990-A
: CURRENT APPLICATION NUMBER: US/09/843,905A
: CURRENT FILING DATE: 2001-04-27
: PRIOR APPLICATION NUMBER: US 60/200,198
: PRIOR FILING DATE: 2000-04-28
: NUMBER OF SEQ ID NOS: 15
: SOFTWARE: PatentIn version 3.1
: SEQ ID NO 15
: LENGTH: 458
: TYPE: PRT
: ORGANISM: Caenorhabditis elegans
US-09-843-905A-15

```

| | | | | |
|-----------------------|------------------|-------------------|------------|-------------|
| Query Match | 39.4%; | Score 903; | DB 2; | Length 458; |
| Best Local Similarity | 43.4%; | Pred. No. 1.8e-8; | | |
| Matches 181; | Conservative 70; | Mismatches 146; | Indels 20; | Gaps 7; |

[illegible]

RESULT 10
US-09-252-991A-25700
Sequence 25700, Application US/09252991A
Patent No. 6551795
GENERAL INFORMATION:
APPLICANT: Marc J. Rubenfield et al.
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
TITLE OF INVENTION: AERGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
FILE REFERENCE: 107196.136
CURRENT APPLICATION NUMBER: US/09/252,991A
CURRENT FILING DATE: 1999-02-18
PRIOR APPLICATION NUMBER: US 60/074,788
PRIOR FILING DATE: 1998-02-18
PRIOR APPLICATION NUMBER: US 60/094,190
PRIOR FILING DATE: 1998-07-27
NUMBER OF SEQ ID NOS: 31142
SEQ ID NO 25700
LENGTH: 710
TYPE: PRT
ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-25700

Query Match 4.5%; Score 104; DB 2; Length 710;
Best Local Similarity 22.3%; Pred. No. 0.1;
Matches 90; Conservative 44; Mismatches 162; Indels 108; Gaps 21;

QY 29 GALEPNDGRGRKSRFALYKRPKANGKSTVHVIPTQASKAIS-----CKGQ 76
DB 264 GLQPDARRRRAAVIGF-RARRGAPPRIVGTMQSANBELTTDIAVIGGAAGI 322
QY 77 HHSITSLRQTVVYETHDKDTPMFOVGRSTESPDPVPTDTTSSQNTDEAOITOSTI 136
DB 323 GVIAASLKRSPGALISLIEPADTHYQPGWTLVGGGAAGQDTA-----RPMGLVPGV 377
QY 137 SRPACRIVCRNEPYTARIPAFGDS--SKNIPL--GEKAKMKNPDGMDGLTNTGV-- 190
DB 378 EMKTRV--RVDPBARLLLEGDSLEYRNLIVCPGLRLA-WERISGLETTLRNVTYS 434
QY 191 -----LVMPHRCG--FTESOPGVWREISVCGDVYTLRETRSAOGRKLVES 235
DB 435 NYRYDLAPYTWELVGRGKALFTGPAMP-----ICAG-----APQKAWYLSC 479
QY 236 E---TNVLDDGSL-IDLCGATILMRPADGLFHTPTQKHLEALROEINARPOCPVGLANT 290
DB 480 DHMLRECVLDDIEVEFDLAGAA-LFGVAD--FVDPMEVYKYSABL-----AFNSNL 529
QY 291 LAPPSIRKEVEKEQWAVYLSGCHVGHNMGRSDTEAN---ERECPMCRVTGPPV 346
DB 530 VKYDGAARKA-----WPEVKADGNTNLAERKDFDLHVVPPQLP 568
QY 347 -----LMLGCEAGF-YVDAGPPTHA---FTPCGHVCESEKSAK 379
DB 569 PTFVAASGLDPAAGWCEVDPATLQHVHGEIFALGDVCGTANAK 612

RESULT 11
US-09-077-098A-6
Sequence 6, Application US/09077098A
Patent No. 6544519
GENERAL INFORMATION:
APPLICANT: TOKUNAGA, Ei-ji
SAKAGUCHI, Masaashi
MATSUO, Kazuo
HAMADA, Fukuaburo
TOKIYOSHI, Sachio
TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS
PARAGALLINARUM AND PROCESS FOR PREPARING THE SAME
NUMBER OF SEQUENCES: 8
CORRESPONDENCE ADDRESS:
ADDRESSEE: BROWDY AND NEIMARK
STREET: 624 Ninth Street, N.W., Suite 300
CITY: Washington

STATE: D.C.
COUNTRY: USA
ZIP: 20001
COMPUTER READABLE FORM:
MEDIUM TYPE: floppy disk
COMPUTER: IBM PC compatible
OPERATING SYSTEM: PC-DOS/MS-DOS
SOFTWARE: Patent in Release #1.0, Version #1.30
CURRENT APPLICATION DATA:
APPLICATION NUMBER: US/09/077,098A
FILING DATE: 19-May-1998
PRIOR APPLICATION DATA:
APPLICATION NUMBER: PCT/JP97/03222
FILING DATE: 12-SEP-1997
APPLICATION NUMBER: JP 27,148/1996
FILING DATE: 19-SEP-1996
ATTORNEY/AGENT INFORMATION:
NAME: KORNBAU, Anne M.
REGISTRATION NUMBER: 25,618
REFERENCE/DOCKET NUMBER: TOKUNAGA-1
TELECOMMUNICATION INFORMATION:
TELEPHONE: 202-628-5197
TELEFAX: 202-737-3528
INFORMATION FOR SEQ ID NO: 6:
SEQUENCE CHARACTERISTICS:
LENGTH: 2042 amino acids
TYPE: amino acid
STRANDEDNESS: single
TOPOLOGY: linear
MOLECULE TYPE: peptide
SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-09-077-098A-6

Query Match 4.4%; Score 100; DB 2; Length 2042;
Best Local Similarity 22.7%; Pred. No. 1.6;
Matches 73; Conservative 45; Mismatches 124; Indels 80; Gaps 18;

QY 53 GKVPSTVHYVISTP--QASKAISCKGQHSIS---YTLRNRQTVVYETHDKD-TDMFOVG 105
DB 1346 GLDDTTLINKNNPADDLNLSSESQKNAITGLVDVVKKTNPSPTVSPSTDSNKKKFTVG 1405
QY 106 RSTESPDPVPTDTTSSQNTDEAOITQS-----TISRPAKIVCD-----R 147
DB 1406 -----VDF--TYTTEGDAYDDKLLTTSKVSYSYTNKLANFTDILSDGSGNAATTA 1457
QY 148 NEPYTARIPAFGD--SKNIPLGEEK-----AAKKNPDG-----HMDGLTNT--GV 190
DB 1458 NDQVGRRLSDGTTISENFTLGSKQYNGSDSLGVWYDDQNGVFKLSLNTALTTLTSLANT 1517
QY 191 LVNHPRGFTESOPGVWREISVCGDVYTLRETRSAOGRKLVESNTVL-----QDGS- 244
DB 1518 FAKLDASNLTDSENKEKMR---TALNVYSKTEVDAREIKSKVTLTDPDGLIFATKQAGSG 1574
QY 245 ---LIDLCGATILMRPADGLFHTPT-----OKHLEALROEINARPOCPVG 287
DB 1575 NNAGID-AGKKISNVADGDI-SPTSGDVYVGRQLYALMOKGIRVYGVDEVSPKTYTAP 1632
QY 288 LNTLAPPSIRKEVEKEQWPA 309
DB 1633 TN--ANPTATTAPTAASQTGMA 1652

RESULT 12
US-10-192-584-6
Sequence 6, Application US/10192584
Patent No. 6919080
GENERAL INFORMATION:
APPLICANT: TOKUNAGA, Ei-ji
SAKAGUCHI, Masaashi
MATSUO, Kazuo
HAMADA, Fukuaburo
TOKIYOSHI, Sachio
TITLE OF INVENTION: NOVEL POLYPEPTIDE FROM HAEMOPHILUS

```

/ PARAGALLINARUM AND PROCESS FOR PREPARING THE SAME
/
/ NUMBER OF SEQUENCES: 8
/ CORRESPONDENCE ADDRESS:
/ ADDRESSEE: BROWDY AND NEIMARK
/ STREET: 624 Ninth Street, N.W., Suite 300
/ CITY: Washington
/ STATE: D.C.
/ COUNTRY: USA
/ ZIP: 20001
/
/ COMPUTER READABLE FORM:
/ MEDIUM TYPE: Floppy disk
/ COMPUTER: IBM PC compatible
/ OPERATING SYSTEM: PC-DOS/MS-DOS
/ SOFTWARE: Patent in Release #1.0, Version #1.30
/ CURRENT APPLICATION DATA:
/ APPLICATION NUMBER: US/10/192,584
/ FILING DATE: 11-Jul-2002
/
/ PRIOR APPLICATION DATA:
/ APPLICATION NUMBER: US/09/077,098
/ FILING DATE: 19-May-1998
/ APPLICATION NUMBER: PCT/JP97/03222
/ FILING DATE: 12-SEP-1997
/ APPLICATION NUMBER: JP 27,148/1996
/ FILING DATE: 19-SEP-1996
/ ATTORNEY/AGENT INFORMATION:
/ NAME: KORNBAU, Anne M.
/ REGISTRATION NUMBER: 25,618
/ REFERENCE/DOCKET NUMBER: TOKUNAGA=1
/ TELECOMMUNICATION INFORMATION:
/ TELEPHONE: 202-628-5197
/ TELEFAX: 202-737-3528
/
/ INFORMATION FOR SEQ ID NO: 6:
/ SEQUENCE CHARACTERISTICS:
/ LENGTH: 2042 amino acids
/ TYPE: amino acid
/ STRANDEDNESS: single
/ TOPOLOGY: linear
/ MOLECULE TYPE: peptide
/ SEQUENCE DESCRIPTION: SEQ ID NO: 6:
US-10-192-584-6

```

```

Query Match 4.4%; Score 100; DB 2; Length 2042;
Best Local Similarity 22.7%; Pred. No. 1.6;
Matches 73; Conservative 45; Mismatches 124; Indels 80; Gaps 18;

QY 53 GVKPSTHYVISTP-QASKAISCKGSHS-----YTSRQTVVVEYTHDKD-TDMQVG 105
DB 1346 GLDPTLTKINNPADQSLNLSSESKNAITGLVDVVKTKNSPIVPEPSTDSNKKQPTVG 1405
QY 106 RSTESPIDFVVTIISGSONTDEAQTOS-----TISRFACRIVCD-----R 147
DB 1406 -----VDF--TDITTEGDATDDKGLTTTSKSVESYVTKLANPSTDILLDGSAGNATTA 1457
QY 148 NEPYTARIFAAGPD-SKKNIFLGEK-----AAKKNPDG-----HMDGLTTN--GV 190
DB 1458 NDGVGKRRLLSDGFTIKSENFLGSKQYNGSDSLGVMVDQNGVFKLSLMTALTTSLANT 1517
QY 191 LVNHPRGFTFESGPGVWREISVCGDYVTLRETRSAQQRGLVSESTNVL-----QDGS- 244
DB 1518 PAKLDASNLTDSDSKKEMR--TALNVYSKTEVDAAEQSKSVTLTPDSGLIFATKQAGSG 1574
QY 245 ---LIDCGATLWRTDGLFHTPT-----QKHIEALQEIINARPOCPVG 287
DB 1575 NNAGID-AGKKKISNVADGDI-SPTSGDVVTGRQLYALMKGKIRVYGDDEVSPYTKQTTPAR 1632
QY 288 LNTLAPPSINRKEVEEKOPWA 309
DB 1633 TN--ANPTATTATPAASSTQGWA 1652

```

RESULT 13
US-09-949-016-6500
; Sequence 6500, Application US/09949016

```

/ Patent No. 6812339
/ GENERAL INFORMATION:
/ APPLICANT: VENTER, J. Craig et al.
/ TITLE OF INVENTION: POLYMORPHISMS IN KNOWN GENES ASSOCIATED
/ TITLE OF INVENTION: WITH HUMAN DISEASE, METHODS OF DETECTION AND USES THEREOF
/ FILE REFERENCE: CL001307
/ CURRENT APPLICATION NUMBER: US/09/949,016
/ CURRENT FILING DATE: 2000-04-14
/ PRIOR APPLICATION NUMBER: 60/241,755
/ PRIOR FILING DATE: 2000-10-20
/ PRIOR APPLICATION NUMBER: 60/237,768
/ PRIOR FILING DATE: 2000-10-03
/ PRIOR APPLICATION NUMBER: 60/231,498
/ PRIOR FILING DATE: 2000-09-08
/ NUMBER OF SEQ ID NOS: 207012
/ SOFTWARE: PasteSeq for Windows Version 4.0
/ SEQ ID NO 6500
/ LENGTH: 717
/ TYPE: PRT
/ ORGANISM: Human
US-09-949-016-6500

```

```

Query Match 4.2%; Score 96; DB 2; Length 717;
Best Local Similarity 22.4%; Pred. No. 0.76;
Matches 71; Conservative 44; Mismatches 110; Indels 92; Gaps 16;

QY 52 NGVKPSTHYVISTPQASKAISCK-----KGSHSYTSLSRNQTVVVEYTHDKDQWFOYG 105
DB 211 DSVTLTNTLHVIGIGFPRAVTCLAFSKNSGNGNLCAVDSDNDHVLVSWMQKEKLAADYK 270
QY 106 RSTES--PIDFVVTD---ISGSONTDEAQTOSTI-----SRFACRIVC-----D 146
DB 271 CSNEAVFADPHRTDNIIVTQGEITSLDDPRRLPLRSKDYSTRKAKVVLCTPSE 330
QY 147 RNEPYTARIFAAGPDSSKNIFL---GEKAKKNPDGMDGLT-----TNGVLVNHMPRG 198
DB 331 NGDTITG-----DSSGNILWKGKTNRI SYAVQGAHGGISPLCMLDGTLV---SGG 380
QY 199 FTESGPGVWREISVCGDYVTLRETRSAQQRGL---VSESTNVLDGSLDLCGAT--- 252
DB 381 GKDR-----KLSWSGNQKURKTEIPPOFGPIFVAAGKGVV-----LIGTTNPF 427
QY 253 LWMRTDGLFHTPTQKRIAL-ROEINARPOCPVGLNTLAPPSINRKEVEEKOPWAYL 311
DB 428 VLQGTISGPTPTTQGTDELMGLAIHASKPQ-----FL 461
QY 312 SCGHVGHYHNW---GHR 325
DB 462 TCGHDKHATLMDAVGHR 478

```

```

RESULT 14
US-09-198-452A-485
; Sequence 485, Application US/09198452A
; Patent No. 6559294
; GENERAL INFORMATION:
; APPLICANT: Grifols, R.
; TITLE OF INVENTION: Chlamydia pneumoniae genomic sequence and polypeptides, fragments
; TITLE OF INVENTION: thereof and uses thereof, in particular for the diagnosis, prever
; FILE REFERENCE: 9710-003-999
; CURRENT APPLICATION NUMBER: US/09/198,452A
; CURRENT FILING DATE: 1998-11-24
; NUMBER OF SEQ ID NOS: 6849
; SEQ ID NO 485
; LENGTH: 492
; TYPE: PRT
; ORGANISM: Chlamydia pneumoniae
US-09-198-452A-485

```

Query Match 4.2%; Score 95.5; DB 2; Length 492;
Best Local Similarity 22.1%; Pred. No. 0.45;
Matches 47; Conservative 25; Mismatches 74; Indels 67; Gaps 9;

```

QY 102 CQVSGSTSPDIPVAVDTIGSQONMDEAOTOSTIRFACRIJCDBN----- 149
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 287 FTIGCADDPBSFLKRDPEVSASVMACTKVT-----LIGALVLDHDDYDLDYMSLOS 340
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 150 PYTARIFAPGDPSSKNITFLGEAKRWKNDGHDGLTNG-----VLVNHPRG 197
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 341 PVAPPI-----AVFKGATVTYTKGFPDGEIATPSHYGYQGKWSYTSRPLLIAPDG 391
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 198 GFTESSQP-----GWRISISVCGDYVTLRETRSQQQRKLYSE-----TNVL 240
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 392 GFPGGPPSSANTLYAVWMSDTLVRSTYLI-----DPRYGEIYSNLSMISFLGNQAFSDIL 447
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

QY 241 QDGSILIDCGATLLWRTADGL-----FHTPTOKH 269
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|
Db 448 QDVLILIDHPLSTI---TAKALCAVYEHTRROGH 477
    |:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|:::|

```

RESULT 15

```

US-10-197-220-65
: Sequence 65, Application US/10197220
: Patent No. 6919187
:
: GENERAL INFORMATION:
:
: APPLICANT: Bhatia, Ajay
:
: APPLICANT: Guderian, Jeff
:
: APPLICANT: Skelky, Yasir A. W.
:
: APPLICANT: Malsbomeuve, Jean-Francois L.
:
: TITLE OF INVENTION: COMPOUNDS AND METHODS FOR TREATMENT
:
: TITLE OF INVENTION: AND DIAGNOSIS OF CHLAMYDIAL INFECTION
:
: FILE REFERENCE: 210121.515C3
:
: CURRENT APPLICATION NUMBER: US/10/197,220
:
: CURRENT FILING DATE: 2002-07-15
:
: NUMBER OF SEQ ID NOS: 175
:
: SEQ ID NO 65
:
: LENGTH: 978
:
: TYPE: prt
:
: ORGANISM: Chlamydia pneumoniae
:
US-10-197-220-65

```

| | | | | |
|-----------------------|-----------------|---------------|-----------|------------|
| Query Match | 4.2% | Score 95.5 | DB 2 | Length 978 |
| Best Local Similarity | 22.1% | Pred. No. 1.4 | | |
| Matches 47 | Conservative 25 | Mismatches 74 | Indels 67 | Gaps 9 |

[illegible]

```
Search completed: March 1, 2006, 21:01:40
Job time : 49 secs
```

This Page Blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2006, 20:53:05 ; Search time 232 Seconds
(without alignments)
1277.250 Million cell updates/sec

Title: US-10-041-030-4

Perfect score: 2290

Sequence: 1 MFSPQGEHCAKPKKPKYK.....ATQLVGEQNCIKLIFGCPID 420

Scoring table: BLOSUM62
Gapop 10.0 , Gapept 0.5

Searched: 2166443 seqs, 705528306 residues

Total number of hits satisfying chosen parameters: 2166443

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%

Listing first 45 summaries

Database : Uniprot 05.80:*
1: uniprot_sprot:*
2: uniprot_trembl:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----------------|---------------------|
| 1 | 2290 | 100.0 | 420 | 1 PELI2_HUMAN | Q9HAE8 homo sapien |
| 2 | 2172.5 | 94.9 | 419 | 1 PELI2_MOUSE | Q8B8E6 mus musculu |
| 3 | 2138.5 | 93.4 | 419 | 2 O6G057_XENLA | Q6G957 xenopus lae |
| 4 | 2129.5 | 93.0 | 419 | 2 O6P4Y5_XENTR | Q6P4Y5 xenopus tro |
| 5 | 2006 | 87.6 | 428 | 2 O6DHJ8_BRARE | O6DHJ8 brachydanio |
| 6 | 1917 | 83.7 | 418 | 1 PELI1_HUMAN | Q96F31 homo sapien |
| 7 | 1917 | 83.7 | 418 | 2 O53126_HUMAN | Q53126 homo sapien |
| 8 | 1912 | 83.5 | 418 | 1 PELI1_MOUSE | Q8C669 mus musculu |
| 9 | 1912 | 83.5 | 418 | 2 O5SRW7_MOUSE | O5SRW7 mus musculu |
| 10 | 1884 | 82.3 | 418 | 2 O7ZXU3_XENLA | O7ZXU3 xenopus lae |
| 11 | 1864 | 81.4 | 405 | 2 O5ZKT7_CHICK | O5ZKT7 gallus gall |
| 12 | 1788 | 78.1 | 428 | 2 O4SGC9_TETNG | O4SGC9 tetradon n |
| 13 | 1647 | 71.9 | 440 | 2 O4SEB1_TETNG | O4SEB1 tetradon n |
| 14 | 1632 | 71.3 | 445 | 1 PELI3_MOUSE | Q8B8E6 mus musculu |
| 15 | 1620 | 70.7 | 469 | 1 PELI3_HUMAN | Q8B8E6 mus musculu |
| 16 | 1493 | 65.2 | 424 | 1 O4SGC7_TETNG | O4SGC7 tetradon n |
| 17 | 1302 | 56.9 | 424 | 1 PELI1_MOUSE | O77237 drosophila |
| 18 | 1279.5 | 55.9 | 389 | 2 O7POM0_ANOGA | O7POM0 anopheles g |
| 19 | 1235 | 53.9 | 455 | 2 O9NDP9_CIOIN | O9NDP9 ciona intes |
| 20 | 1100 | 48.0 | 441 | 2 O4RLT6_TETNG | O4RLT6 tetradon n |
| 21 | 910 | 39.7 | 458 | 2 O22967_CABEL | O22967 caenorhabdi |
| 22 | 903 | 39.4 | 458 | 2 O9NAP3_CABEL | O9NAP3 caenorhabdi |
| 23 | 883.5 | 38.6 | 450 | 2 O61KK0_CABER | O61KK0 caenorhabdi |
| 24 | 602 | 26.3 | 121 | 2 O562B8_RAT | O562B8 rattus norv |
| 25 | 446.5 | 19.5 | 146 | 2 O6Q430_DROBU | O6Q430 drosophila |
| 26 | 277 | 12.1 | 49 | 2 O659D8_HUMAN | O659D8 homo sapien |
| 27 | 264 | 11.5 | 75 | 2 O5ZM03_CHICK | O5ZM03 gallus gall |
| 28 | 212 | 9.3 | 57 | 2 O4RAV6_TETNG | O4RAV6 tetradon n |
| 29 | 122.5 | 5.3 | 215 | 2 O9YVJ8_MSEPP | O9YVJ8 melanoplus n |
| 30 | 119 | 5.2 | 340 | 2 O8SVY5_ENCCU | O8SVY5 encephalito |
| 31 | 118 | 5.2 | 1358 | 2 O7QC06_ANOGA | O7QC06 anopheles g |

ALIGNMENTS

| RESULT 1 | PELI2_HUMAN | STANDARD; | PRT; | 420 AA. |
|----------|---|-----------|------|---------|
| AC | Q9HAE8; | | | |
| DT | 23-MAR-2004 (Rel. 43, Created) | | | |
| DT | 29-MAR-2004 (Rel. 43, Last sequence update) | | | |
| DT | 10-MAY-2005 (Rel. 47, Last annotation update) | | | |
| DE | Pellino protein homolog 2 (Pellino 2). | | | |
| GN | Name=PELI2; | | | |
| OS | Homo sapiens (Human). | | | |
| OC | Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; | | | |
| OC | Mammalia; Eutheria; Barchontoglires; Primates; Catarrhini; Homidae; | | | |
| OC | Homo. | | | |
| OX | NCBI_TaxID=9606; | | | |
| RN | [1] | | | |
| RP | NUCLEOTIDE SEQUENCE. | | | |
| RX | MEDLINE=21203570; PubMed=11306823; | | | |
| RA | Resch K., Jockusch H., Schmitt-John T.; | | | |
| RT | "Assignment of homologous genes, Pel11/Pel11 and Pel12/Pel12, for the | | | |
| RT | Pelle adaptor protein Pellino to mouse chromosomes 11 and 14 and human | | | |
| RT | chromosomes 2p13.3 and 14q21, respectively, by physical and radiation | | | |
| RT | hybrid mapping." | | | |
| RL | Cytogenet. Cell Genet. 92:172-174(2001). | | | |
| RN | [2] | | | |
| RP | NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA]. | | | |
| RC | TISSUE=Lung; | | | |
| RX | MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899; | | | |
| RA | Strausberg R.L., Pelinoid E.A., Grouse L.H., Derge J.G., | | | |
| RA | Klausner R.D., Collins P.S., Wagner L., Shennan C.M., Schuler G.D., | | | |
| RA | Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K., | | | |
| RA | Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F., | | | |
| RA | Diatchenko L., Matkunas K., Farmer A.A., Rubin G.M., Hong L., | | | |
| RA | Scapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E., | | | |
| RA | Brownstein M.J., Udell T.B., Tothiyuki S., Carninci P., Prange C., | | | |
| RA | Raba S.S., Lomeliello N.A., Peters G.J., Abramson R.D., Mullaby S.J., | | | |
| RA | Bosak S.A., McKernan P.J., McKernan K.U., Malek J.A., Gamarane P.H., | | | |
| RA | Richardson S., Worley K.C., Hale S., Garcia A.M., Gay L.D., Hulyk S.W., | | | |
| RA | Villalongo D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., | | | |
| RA | Fahy J., Hailton B., Ketterman M., Madan A., Rodriguez P., Sanchez A., | | | |
| RA | Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G., | | | |
| RA | Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., | | | |
| RA | Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M., | | | |
| RA | Butterfield Y.S.N., Krzywnicki M.T., Skalska U., Smalins D.E., | | | |
| RA | Schneerch A., Schein J.E., Jones S.J.M., Marra M.A.; | | | |
| RT | "Generation and initial analysis of more than 15,000 full-length human | | | |
| RT | and mouse cDNA sequences." | | | |
| RL | Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002). | | | |
| RN | [3] | | | |
| RP | FUNCTION AND INTERACTIONS WITH TRAP6 AND MAP3K7. | | | |
| RX | MEDLINE=22689054; PubMed=12804775; DOI=10.1016/S0014-5793(03)00533-7; | | | |
| RA | Jensen L.E., Whitehead A.S.; | | | |
| RT | "Pellino2 activates the mitogen activated protein kinase pathway." | | | |
| RL | FEBS Lett. 545:199-202(2003). | | | |
| RN | [4] | | | |

| | | | | | | |
|----|-------|-----|------|---|--------------|--------------------|
| 32 | 116.5 | 5.1 | 1084 | 2 | Q6NK13_CORDI | Q6NK13 corynebacte |
| 33 | 113 | 4.9 | 804 | 2 | O7TPT4_MOUSE | O7TPT4 mus musculu |
| 34 | 113 | 4.9 | 818 | 2 | O8CC59_MOUSE | O8CC59 mus musculu |
| 35 | 113 | 4.9 | 818 | 2 | O9DBC8_MOUSE | O9DBC8 mus musculu |
| 36 | 105.5 | 4.6 | 1060 | 2 | O59GK8_HUMAN | O59GK8 homo sapien |
| 37 | 103.5 | 4.5 | 1591 | 1 | TIAM1_MOUSE | O6B10 mus musculu |
| 38 | 103 | 4.5 | 675 | 2 | O7QX14_GIALA | O7QX14 giardia lam |
| 39 | 102.5 | 4.5 | 1857 | 2 | O8TW75_METAC | O8TW75 methanocarc |
| 40 | 102 | 4.5 | 620 | 2 | O96DH7_HUMAN | O96DH7 homo sapien |
| 41 | 102 | 4.5 | 689 | 2 | O9URU4_MOUSE | O9URU4 mus musculu |
| 42 | 102 | 4.5 | 710 | 2 | O6PB38_MOUSE | O6PB38 mus musculu |
| 43 | 102 | 4.5 | 711 | 2 | O8BQC8_MOUSE | O8BQC8 mus musculu |
| 44 | 102 | 4.5 | 815 | 2 | O8WVL6_HUMAN | O8WVL6 homo sapien |
| 45 | 101.5 | 4.4 | 900 | 2 | Q6PID6_MOUSE | Q6PID6 mus musculu |

RP FUNCTION, PHOSPHORYLATION, AND INTERACTIONS WITH IRAK1 AND IRAK4.
 RX MEDLINE=22744764; PubMed=12860405; DOI=10.1016/S0014-5793(03)00697-5;
 RA Strelow A., Kollwe C., Wesche H.;
 RT "Characterization of Pellino2, a substrate of IRAK1 and IRAK4.";
 RL FEBS Lett. 547:157-161(2003).
 CC -1- FUNCTION: Scaffold protein which probably links Toll-like
 CC receptors (TLRs) to basic cellular processes via its interaction
 CC with the complex containing IRAK kinases and TRAF6. Can activate
 CC the MAP (mitogen activated protein) kinase pathway leading to
 CC activation of ERK1. Not required for NF-kappa-B activation.
 CC -1- SUBUNIT: Interacts with TRAF6, IRAK1, IRAK4 and MAP3K7.
 CC -1- INTERACTION:
 CC P51617:IRAK1; NExp=1; Intact=EBI-448407, EBI-358664;
 CC Q9NMW3:IRAK4; NExp=1; Intact=EBI-448407, EBI-448378;
 CC -1- PTM: Phosphorylated by IRAK1 and IRAK4.
 CC -1- SIMILARITY: Belongs to the pellino family.
 CC -----
 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC -----
 CC EMBL; AF302502; AAG15390.1; -, mRNA.
 CC EMBL; BC009476; AAH09476.1; -, mRNA.
 CC Inactive; Q9HAT8; -.
 CC DR Ensembl; ENSG00000139946; Homo sapiens.
 CC DR HGNC; HGNC:8829; PELI2.
 CC DR H-InvDB; HIK0011690; -.
 CC DR InterPro; IPR006800; Pellino.
 CC DR PANTHER; PTHR12098; Pellino; 1.
 CC DR Pfam; PF04710; Pellino; 1.
 CC DR KEGG; K04710; Pellino; 1.
 CC KM Phosphorylation.
 CC SO SEQUENCE 420 AA; 46435 MW; 2FC5661C13BC11A CRC64;
 CC -----
 CC Query Match 100.0%; Score 2290; DB 1; Length 420;
 CC Best Local Similarity 100.0%; Pred. No. 1.5e-182;
 CC Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 CC -----
 CC QY 1 MESPQGEHCAPKPEPKYKELVVLGYNGALPNDRRRRSRPALKYGRPKANGKYSTYH 60
 CC Db 1 MESPQGEHCAPKPEPKYKELVVLGYNGALPNDRRRRSRPALKYGRPKANGKYSTYH 60
 CC QY 61 VISTPQASKAISKQGHISYTLSSRNQTVVEYHDTQTMFOYGRSTESPDPVVDTI 120
 CC Db 61 VISTPQASKAISKQGHISYTLSSRNQTVVEYHDTQTMFOYGRSTESPDPVVDTI 120
 CC QY 121 SSSQNTDEAQITOSTISRFACRIYCDNRPYTAIFAAGFDSKNIFLGEKAAKXNDG 180
 CC Db 121 SSSQNTDEAQITOSTISRFACRIYCDNRPYTAIFAAGFDSKNIFLGEKAAKXNDG 180
 CC QY 181 HMGGLTNGVLNHPRGFTPEESQGVWRBISVCGDVYTLRFRSSAQQRKLVSESNVL 240
 CC Db 181 HMGGLTNGVLNHPRGFTPEESQGVWRBISVCGDVYTLRFRSSAQQRKLVSESNVL 240
 CC QY 241 QDSGLIDLCGATLMLRTADGLFHTPTQKHIEALROEINAAPOCPVGLNTLAPSTIRKE 300
 CC Db 241 QDSGLIDLCGATLMLRTADGLFHTPTQKHIEALROEINAAPOCPVGLNTLAPSTIRKE 300
 CC QY 301 VVEEKQWAVLSCGHVGYNHWGRSDTEANERECPCMCRTVGPVPLMLGCEAGFYVDAG 360
 CC Db 301 VVEEKQWAVLSCGHVGYNHWGRSDTEANERECPCMCRTVGPVPLMLGCEAGFYVDAG 360
 CC QY 361 PPTHAFPPCGHVSEKSAKXWSQIPLPHGTHAFPAAPFCATQLVGEQNCIKLIFQCPID 420
 CC Db 361 PPTHAFPPCGHVSEKSAKXWSQIPLPHGTHAFPAAPFCATQLVGEQNCIKLIFQCPID 420
 CC -----
 CC RESULT 2
 CC PELI2_MOUSE STANDARD; PRT; 419 AA.
 CC ID PELI2_MOUSE
 CC AC Q8BST6; Q8CAF2; Q8CC65; Q8R2X4; Q8R9J7;
 CC DT 29-MAR-2004 (Rel. 43, Created)

DT 29-MAR-2004 (Rel. 43, last sequence update)
 DE 10-MAY-2005 (Rel. 47, last annotation update)
 DE Pellino protein homolog 2 (Pellino 2).
 GN Name=Peli2;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Muridae; Murinae; Mus.
 ON NCBI_TaxID=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE (ISOFORM 1).
 RC STRAIN=C57BL/6J;
 RX MEDLINE=21203570; PubMed=11306823;
 RA Resch K., Jockusch H., Schmitt-John T.;
 RT "Assignment of homologous genes, Peli1/PELI1 and Peli2/PELI2, for the
 RT Peli adaptor protein Pellino to mouse chromosomes 11 and 14 and human
 RT chromosomes 2p13.3 and 14q21, respectively, by physical and radiation
 RT hybrid mapping.";
 RL Cytogenet. Cell Genet. 92:172-174(2001).
 RN [2]
 RP NUCLEOTIDE SEQUENCE (LARGE SCALE MRNA) (ISOFORMS 1 AND 3).
 RC STRAIN=C57BL/6J; TISSUE=Cerebellum, Epididymis, and Pituitary;
 RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;
 RA Okazaki Y., Furuno M., Kasukawa T., Adachi Y., Bono H., Kondo S.,
 RA Nkaido I., Otsu N., Saito R., Suzuki H., Yamana I., Kiyosawa H.,
 RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,
 RA Baldarelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,
 RA Schirini L.M., Kampin A., Matsuda H., Batalov S., Beisel K.W.,
 RA Blake U.A., Brad D., Brusic V., Chotila C., Corbani L.E., Cousins S.,
 RA Dalla E., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S.,
 RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,
 RA Grimmond S., Gustincich S., Hirokawa N., Jackson J.J., Jarvis E.D.,
 RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,
 RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,
 RA Maglocz D.R., Maltais L., Marchionni L., McKenna H., Miki H.,
 RA Nagashima T., Numata K., Okido T., Pavan W.U., Petosa G., Pesole G.,
 RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramchandran S.,
 RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,
 RA Sandelin A., Schneider C., Sempke C.A., Setou M., Shimada K.,
 RA Sultana R., Takenawa Y., Taylor M.S., Teasdale R.D., Tomita M.,
 RA Verardo R., Wagner L., Wahlstedt C., Wang Y., Watanabe Y., Wells C.,
 RA Wilming L.G., Wyshaw-Borls A., Yanagisawa M., Yang I., Yang L.,
 RA Yuan Z., Zavalan M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,
 RA Hirose-Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,
 RA Shiraki T., Waki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
 RA Hara A., Hashizume W., Inotani K., Ishii Y., Itoh M., Kagawa I.,
 RA Miyazaki A., Sakai K., Sasaki D., Shidaka K., Shingawa A.,
 RA Yasunishi A., Yoshino Y., Waterston R., Lander E.S., Rogers J.,
 RA Birney E., Hayashizaki Y.;
 RT "Analysis of the mouse transcriptome based on functional annotation of
 RT 60,770 full-length cDNAs.";
 RL Nature 420:563-573(2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE (LARGE SCALE MRNA) (ISOFORM 2).
 RC TISSUE=Mammary tumor;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strassberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shellen C.M., Schlier G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.P., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heish F.,
 RA Dackiwko L., Marusika K., Farmer A.A., Rubin G.M., Hong L.,
 RA Brownstein M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hultky S.W.,
 RA Villalon D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton B., Kerteman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butterfield Y.S.N., Krzywicki M.I., Skalska U., Shailus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.;

| | |
|----|--|
| RT | "Generation and initial analysis of more than 15,000 full-length human |
| RL | and mouse cDNA sequences".; |
| RN | Proc. Natl. Acad. Sci. U.S.A. 99:16889-16903(2002). |
| RP | [4] |
| RX | FUNCTION, TISSUE SPECIFICITY, AND INTERACTION WITH IRAK1. |
| RA | MEDLINE=22257640; PubMed=12370331; |
| RA | Yu K.-Y., Kwon H.-J., Norman D.A.M., Vig B., Geobl M.G., |
| RA | Harrington M.A.; |
| RT | "Mouse pellino-2 modulates IL-1 and lipopolysaccharide signaling."; |
| RL | J. Immunol. 169:4075-4078(2002). |
| CC | -I- FUNCTION: Scaffold protein which probably links Toll-like |
| CC | receptors (TLRs) to basic cellular processes via its interaction |
| CC | with the complex containing IRAK kinases and TRAF6. Can activate |
| CC | the MAP (mitogen activated protein) kinase pathway leading to |
| CC | activation of ELK1. Not required for NF-kappa-B activation. |
| CC | -I- SUBUNIT: Interacts with TRAF6, IRAK4 and MAP3K7 (By similarity). |
| CC | -I- INTERACTS: Interacts with IRAK1. |
| CC | -I- INTERACTION: |
| CC | OS2406:Irk4; NBExp=1; IntAct=Ebi-448554, Ebi-448533; |
| CC | -I- ALTERNATIVE PRODUCTS: |
| CC | Event=Alternative splicing; Named isoforms=3; |
| CC | Name=1; |
| CC | IsoId=Q8BST6-1; Sequence=Displayed; |
| CC | Name=2; |
| CC | IsoId=Q8BST6-2; Sequence=VSP_008636; |
| CC | Note=No experimental confirmation available; |
| CC | Name=3; |
| CC | IsoId=Q8BST6-3; Sequence=VSP_008637, VSP_008638; |
| CC | Note=No experimental confirmation available; |
| CC | -I- TISSUE SPECIFICITY: Widely expressed both in embryos and adult. |
| CC | -I- Weakly or not expressed in spleen and thymus. |
| CC | -I- PTM: Phosphorylated by IRAK1 and IRAK4 (by similarity). |
| CC | -I- SIMILARITY: Belongs to the pellino family. |
| CC | ----- |
| CC | This Swiss-Prot entry is copyright. It is produced through a collaboration |
| CC | between the Swiss Institute of Bioinformatics and the EMBL outstation - |
| CC | the European Bioinformatics Institute. There are no restrictions on its |
| CC | use as long as its content is in no way modified and this statement is not |
| CC | removed. |
| CC | ----- |
| DR | EMBL; AF302504; AAG5392.1; -; mRNA. |
| DR | EMBL; AK030564; BAC27024.1; -; mRNA. |
| DR | EMBL; AK033815; BAC28485.1; -; mRNA. |
| DR | EMBL; AK082342; BAC38472.1; -; mRNA. |
| DR | EMBL; BC027062; AAH27062.1; -; mRNA. |
| DR | INLACC; Q8BST6; - |
| DR | Ensembl; ENSMGCG0000021846; Mus musculus. |
| DR | MGI; MGI:1891445; Pel12. |
| DR | GO; GO:0005515; P:protein binding; IPI. |
| DR | InterPro; IPRO06800; Pellino. |
| DR | PANTHER; PTHR12098; Pellino. |
| DR | Pfam; PF04710; Pellino; 1. |
| KW | Alternative splicing; Phosphorylation. |
| FT | VASPLIC |
| FT | 1 |
| FT | 100 |
| FT | Missing (in isoform 2). |
| FT | /ftid=vsp_008636. |
| FT | AISGSHGISITYSRSQVVVEYTHRDOTMPGVNSTES |
| FT | PIDFVPTDVSQGNEADQIOSTISRACIVCRNPYT |
| FT | ARIF -> LPAKHYNENADESLALPLTGLTGCSO |
| FT | RRRYDDPAREGRGASGNQVVAQAFLICPLSYTYKQBOIR |
| FT | CLKIKPFSCMS (in isoform 3). |
| FT | /ftid=vsp_008637. |
| FT | Missing (in isoform 3). |
| FT | /ftid=vsp_008638. |
| FT | E -> K (in Ref. 2; BAC27024). |
| FT | A -> T (in Ref. 2; BAC27024). |
| FT | R -> G (in Ref. 2; BAC38472). |
| FT | RR -> KK (in Ref. 1). |
| FT | A -> T (in Ref. 1). |
| FT | S -> H (in Ref. 1). |
| FT | T -> A (in Ref. 3). |
| FT | R -> M (in Ref. 1). |
| FT | A -> V (in Ref. 1). |
| FT | SEQUENCE |
| FT | 419 AA; 46272 MW; 786C92C28C3BD0CB CRC64; |
| QO | |

| Query Match | 94.9% | Score 2172.5 | DB 1 | Length 419 |
|-----------------------|---|--|----------|------------|
| Best Local Similarity | 95.0% | Pred. No 196-173 | | |
| Matches 399 | Conservative 10 | Mismatches 10 | Indels 1 | Gaps 1 |
| Qy | 1 | MFSPDSEHCAPNKEPVYKGLVVLGYNGALPNGDRGRKSRFALYRRPKANGVKSPTVH | 60 | |
| Db | 1 | MFSPDSEHCAPNKEPVYKRELVLGYNGALPNGDRGRKSRFALYKRTVASGVSPSTIH | 60 | |
| Qy | 61 | VISTPQASKAISCKGCHSISTYLSRNOTVVEYTHDDTQMPQVGRSTESPIDPVYDTI | 120 | |
| Db | 61 | WVSTPQASKAISCKGCHSISTYLSRQTVVVEYTHDDTQMPQVGRSTESPIDPVYDTV | 120 | |
| Qy | 121 | SGSONTDEAOITQSTIRPACRIYCDNREPYTAIIPAAGPSSKNIFGEEAAAKKNDG | 180 | |
| Db | 121 | SGGQNEQ-AQITQSTIRPACRIYCDNREPTAIIPAAGPSSKNIFGEEAAAKKNDG | 179 | |
| Qy | 181 | HMDGLTTNGVLMHPRGFTFEEOPGWREISVCGDVTTLRETRSAOQKGLVSEETVL | 240 | |
| Db | 180 | HMDGLTTNGVLMHPRGFTFEEOPGWREISVCGDVTTLRETRSAOQKGLVSEETVL | 239 | |
| Qy | 241 | QDGLLIDLCGATLLMRTADGLFHTPTQKHEALBOEINAAAPQCPVGLNTLAFPSINKE | 300 | |
| Db | 240 | QDGLLIDLCGATLLMRTADGLFHAPTQKHEALBOEINAAAPQCPVGLNTLAFPSINKE | 299 | |
| Qy | 301 | VVEKQPMAYISCGHVGYNHWGRSDTEANERECPCMKRTVGPVPLMLGCEAGFYVDAG | 360 | |
| Db | 300 | VVEKQPMAYISCGHVGYNHWGRSDTEANERECPCMKRTVGPVPLMLGCEAGFYVDAG | 359 | |
| Qy | 361 | PPTAFTPCGHCVCSEKSAKTSQIPLPHGTHAFHAPACFCATQVGEQNCIKLIFQGPID | 420 | |
| Db | 360 | PPTAFTPCGHCVCSEKSAKTSQIPLPHGTHAFHAPACFCATQVGEQNCIKLIFQGPVD | 419 | |
| RESULT 3 | | | | |
| ID | Q6G057_XENTIA | PRELIMINARY | PRT | 419 AA. |
| AC | Q6G057 | | | |
| DT | 05-JUL-2004 | (TREMBLrel. 27, Created) | | |
| DT | 05-JUL-2004 | (TREMBLrel. 27, Last sequence update) | | |
| DT | 05-JUL-2004 | (TREMBLrel. 27, Last annotation update) | | |
| DE | MG80329 | protein. | | |
| GN | Name=MG80329 | | | |
| OS | Xenopus laevis | (African clawed frog). | | |
| OC | Eukaryota; Metazoa; Chordata; Craniota; Vertebrata; Euteleostomi; | | | |
| CC | Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae; | | | |
| CC | Xenopodinae; Xenopus; Xenopus. | | | |
| OX | NCBI_TaxID=8355; | | | |
| ON | [1] | | | |
| RP | NUCLEOTIDE SEQUENCE. | | | |
| RC | TISSUE=Ovary; | | | |
| RX | MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899; | | | |
| RA | Klauber R.L., Reingold E.A., Grouse L.H., Derge J.G., | | | |
| RA | Klauber R.D., Collins F.S., Wagner L., Shemen C.M., Schuler G.D., | | | |
| RA | Alschul S.F., Zeeberg B., Butow K.H., Schaefer C.F., Bhat N.K., | | | |
| RA | Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F., | | | |
| RA | Datchenko L., Marusina B., Farmer A.A., Rubin G.M., Hong L., | | | |
| RA | Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schetz T.E., | | | |
| RA | Brownstein M.J., Usdin T.B., Toshiyuki S., Carninci P., Prange C., | | | |
| RA | Rana S.S., Loquellano N.A., Peters G.J., Adamson R.D., Mullany S.J., | | | |
| RA | Bosak S.A., McKean P.J., McKernan K.J., Malek J.A., Gunaratne P.H., | | | |
| RA | Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W., | | | |
| RA | Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A., | | | |
| RA | Watley J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A., | | | |
| RA | Whiting M., Madan A., Young A.C., Shcherbina Y., Bouffard G.G., | | | |
| RA | Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C., | | | |
| RA | Butterfield A.C., Grimwood J., Schmutz J., Myers R.M., | | | |
| RA | Butterfield J.S.N., Kiryakin M.I., Skalski U., Smallus D.E., | | | |
| RA | Scherech A., Schein J.E., Jones S.J.M., Marra M.A., | | | |
| RT | "Generation and initial analysis of more than 15,000 full-length human | | | |
| RL | and mouse cDNA sequences". | | | |
| RN | Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903 (2002). | | | |
| | [2] | | | |

RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Ovary;
 RX MEDLINE=2234113; PubMed=12454917; DOI=10.1002/dvdy.10174;
 RA Klein S.L., Struhsberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.,
 RA "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT Initiative";
 RL Dev. Dyn. 225:384-391 (2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Ovary;
 RA Klein S., Gerhard D.S.,
 RA Submitted (JUN-2004) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC072891; AA472891.1; -; mRNA.
 DR Interpro: IPR006800; Pellino.
 DR Pfam: PF04710; Pellino; 1.
 DR SEQUENCE 419 AA; 46372 MW; 55AC2FE7953D6145 CRC64;
 SO
 Query Match 93.4%; Score 2138.5; DB 2; Length 419;
 Best Local Similarity 91.2%; Pred. No. 6.6e-170;
 Matches 383; Conservative 24; Mismatches 12; Indels 1; Gaps 1;
 QY 1 MFSPGQEHKCAPKPEPVYKGLVVLGYNGALPNDGRGRKSRFPALYKRPKANGVKSTVH 60
 DB 1 MFSPGQEHKCAPKPEPVYKGLVVLGYNGCLPNDGRGRKSRFPALYKRPKANGVKSTVH 60
 QY 61 VISTPQASKAISSCKGHSISYTLSSRNQTVVVEYTHDXTDMFQVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQASKAISSCKGHSISYTLSSRNQTVVVEYTHDXTDMFQVGRSTESPIDFVVTDTI 120
 QY 121 SSGQNTDEAQTOSTISRACRIVCDRNPPTARIFAAGFSSKNIFLGEKAAKMPDG 180
 DB 121 SGNQNT-DETQITQSTISRACRIVCDRNPPTARIFAAGFSSKNIFLGEKAAKMPDG 179
 QY 181 HMDGLTNGVLVWHPRGFTBESOPGWREISVCGDVYTLRETRSAQGRKLVESNTVL 240
 DB 180 HMDGLTNGVLVWHPRGFTBESOPGWREISVCGDVYTLRETRSAQGRKLVESNTVL 239
 QY 241 QDGLSLDLCGATLMTADGLFTPTQKHIALROEINARPCQVGLNTLAFPSINRKE 300
 DB 240 QDGLSLDLCGATLMTADGLFTPTQKHIALROEINARPCQVGLNTLAFPSINRKE 299
 QY 301 VVEKOPWAVLSCGHVGHYHNGHRSDEANERECPCRTGVPVPLMLGCEAGFYVDAG 360
 DB 300 VVEKOPWAVLSCGHVGHYHNGHRSDEANERECPCRTGVPVPLMLGCEAGFYVDAG 359
 QY 361 PPTHAFPPCGHVCSKSAKYSQIPLPHGTHAFHACPCATQVGEONCICKIFGCPID 420
 DB 360 PPTHAFPPCGHVCSKSAKYSQIPLPHGTHAFHACPCATQVGEONCICKIFGCPID 419
 RESULT 4
 Q6P4Y5_XENTR PRELIMINARY; PRT; 419 AA.
 ID Q6P4Y5_XENTR PRELIMINARY;
 AC Q6P4Y5;
 DT 05-JUL-2004 (TREMBlrel. 27, Created)
 DT 05-JUL-2004 (TREMBlrel. 27, Last sequence update)
 DT 05-JUL-2004 (TREMBlrel. 27, Last annotation update)
 DE Hypothetical protein WGC75650.
 GN Name=WGC75650;
 OS Xenopus tropicalis (Western clawed frog) (Silurana tropicalis).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipidoidea; Pipidae;
 OC Xenopodidae; Xenopus; Silurana.
 OC NCBI_TaxID=8364;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=22386257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Struhsberg R.L., Reingold E.A., Grouse L.H., Derge J.G.,
 RA Krausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhac N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh F.,

RA Diachenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Schaefer T.E.,
 RA Brownstein M.J., Ustin T.B., Tohbiyuki S., Carninci P., Prange C.,
 RA Raha S.S., Loggellano N.A., Peters G.J., Abramson R.D., Mollahy S.J.,
 RA Bobak S.A., McSwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Huiyik S.W.,
 RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Ketterman M., Madan A., Rodrigues S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko V., Bouffard G.G.,
 RA Blakeley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.W.,
 RA Butlerfield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences";
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RA Klein S., Gerhard D.S.,
 RA Submitted (DEC-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC063200; AA463200.1; -; mRNA.
 DR Interpro: IPR006800; Pellino.
 DR Pfam: PF04710; Pellino; 1.
 DR Hypothetical protein.
 DR SEQUENCE 419 AA; 46328 MW; 042A50786CE69D75 CRC64;
 SO
 Query Match 93.0%; Score 2129.5; DB 2; Length 419;
 Best Local Similarity 91.4%; Pred. No. 3.7e-169;
 Matches 384; Conservative 22; Mismatches 13; Indels 1; Gaps 1;
 QY 1 MFSPGQEHKCAPKPEPVYKGLVVLGYNGALPNDGRGRKSRFPALYKRPKANGVKSTVH 60
 DB 1 MFSPGQEHKCAPKPEPVYKGLVVLGYNGCLPNDGRGRKSRFPALYKRPKANGVKSTVH 60
 QY 61 VISTPQASKAISSCKGHSISYTLSSRNQTVVVEYTHDXTDMFQVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQASKAISSCKGHSISYTLSSRNQTVVVEYTHDXTDMFQVGRSTESPIDFVVTDTI 120
 QY 121 SSGQNTDEAQTOSTISRACRIVCDRNPPTARIFAAGFSSKNIFLGEKAAKMPDG 180
 DB 121 SGNQNT-DETQITQSTISRACRIVCDRNPPTARIFAAGFSSKNIFLGEKAAKMPDG 179
 QY 181 HMDGLTNGVLVWHPRGFTBESOPGWREISVCGDVYTLRETRSAQGRKLVESNTVL 240
 DB 180 HMDGLTNGVLVWHPRGFTBESOPGWREISVCGDVYTLRETRSAQGRKLVESNTVL 239
 QY 241 QDGLSLDLCGATLMTADGLFTPTQKHIALROEINARPCQVGLNTLAFPSINRKE 300
 DB 240 QDGLSLDLCGATLMTADGLFTPTQKHIALROEINARPCQVGLNTLAFPSINRKE 299
 QY 301 VVEKOPWAVLSCGHVGHYHNGHRSDEANERECPCRTGVPVPLMLGCEAGFYVDAG 360
 DB 300 VVEKOPWAVLSCGHVGHYHNGHRSDEANERECPCRTGVPVPLMLGCEAGFYVDAG 359
 QY 361 PPTHAFPPCGHVCSKSAKYSQIPLPHGTHAFHACPCATQVGEONCICKIFGCPID 420
 DB 360 PPTHAFPPCGHVCSKSAKYSQIPLPHGTHAFHACPCATQVGEONCICKIFGCPID 419
 RESULT 5
 Q6DHU8_BRARE PRELIMINARY; PRT; 428 AA.
 ID Q6DHU8_BRARE PRELIMINARY;
 AC Q6DHU8;
 DT 25-OCT-2004 (TREMBlrel. 28, Created)
 DT 25-OCT-2004 (TREMBlrel. 28, Last sequence update)
 DT 25-OCT-2004 (TREMBlrel. 28, Last annotation update)
 DE Pellino homolog 2.
 GN Name=pell2;
 OS Brachydanio rerio (Zebrafish) (Danio rerio).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Actinopterygii; Neopterygii; Teleostei; Ostariophysi; Cypriniformes;
 OC Cyprinidae; Danio.

NCBI_TaxID=7955;
[1]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner C.M., Shenmen C.M., Bhat N.K.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Shat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshitsuki S., Carninci P., Prange C.,
RA Raha S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richardson S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schmech A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[2]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Whole;
RG NIH MGC Project;
RL Submitted (JUL-2004) to the EMBL/GenBank/DBJ databases.
DR EMBL: BC075973; AAH75973.1; -; mRNA.
DR ZFIN: ZDB-GENE-040718-360; zgc:92268.
DR InterPro: IPR006800; Pellino.
DR Pfam: PF04710; Pellino; 1.
SQ SEQUENCE 428 AA; 47287 MW; B33890B49BF4366A CRC64;

Query Match 87.6%; Score 2006; DB 2; Length 428;
Best Local Similarity 84.3%; Pred. No. 8.1e-159;
Matches 361; Conservative 28; Mismatches 31; Indels 8; Gaps 2;

1 MFSPGQGEHCAPNPKPKYKGLVVLGYNGALPNDGRGRKSRFALYKRPANGVSPSTVH 60
1 MFSPGQGEHCAPNPKPKYKGLVVLGYNGALPNDGRGRKSRFALYKRPANGVSPSTVH 60
61 VISTPQASKAISCKGHSISYTLISRNQTVVVEYTHDKDTMPQVGRSTESPIDFVDTT 120
61 VISTPQASKAISCKGHSISYTLISRNQTVVVEYTHDKDTMPQVGRSTESPIDFVDTT 120
121 SGGQNDDEAQTOSTIRFACRIYCDNREPTAIPAAAGDSSKNIFLGEKAAKMKPPDG 180
121 SGGQNDDEAQTOSTIRFACRIYCDNREPTAIPAAAGDSSKNIFLGEKAAKMKPPDG 180
121 SGGQNDDEAQTOSTIRFACRIYCDNREPTAIPAAAGDSSKNIFLGEKAAKMKPPDG 180
121 SGGQNDDEAQTOSTIRFACRIYCDNREPTAIPAAAGDSSKNIFLGEKAAKMKPPDG 180
181 HMDGLTNGVLMVMPRGGFTEESQPGVWRREISVCGDVYTLRETSAAORGKLVESFTNVL 240
181 HMDGLTNGVLMVMPRGGFTEESQPGVWRREISVCGDVYTLRETSAAORGKLVESFTNVL 240
181 HMDGLTNGVLMVMPRGGFTEESQPGVWRREISVCGDVYTLRETSAAORGKLVESFTNVL 240
241 QDSGLIDLCGATILMRTADGLFHTPTQKHLEALROEINARPOCPVGLANTLAPPSINR-- 298
241 QDSGLIDLCGATILMRTADGLFHTPTQKHLEALROEINARPOCPVGLANTLAPPSINR-- 298
241 QDSGLIDLCGATILMRTADGLFHTPTQKHLEALROEINARPOCPVGLANTLAPPSINR-- 298
241 QDSGLIDLCGATILMRTADGLFHTPTQKHLEALROEINARPOCPVGLANTLAPPSINR-- 298
299 -----KEVSEKQPMATLSCGHVGNNGHRSSTEAN-ERECMCTGTPYPLMTGCE 352
299 -----KEVSEKQPMATLSCGHVGNNGHRSSTEAN-ERECMCTGTPYPLMTGCE 352
301 RALSNVPAQEDKDPWVYVLAGCHVGYHDGHSRSEENARRECMCAVGPVPLMTGCE 360
301 RALSNVPAQEDKDPWVYVLAGCHVGYHDGHSRSEENARRECMCAVGPVPLMTGCE 360
353 AGFYVADGPPTHTATPCGHVCESEKAYKSOIPLPHGTHAFHACPCATQOLVCEONCIK 412
353 AGFYVADGPPTHTATPCGHVCESEKAYKSOIPLPHGTHAFHACPCATQOLVCEONCIK 412
361 PAFIVDVGAFTHVVPVPCGHVCESEKAYKSOIPLPHGTHAFHACPCATQOLVCEONCIK 420
361 PAFIVDVGAFTHVVPVPCGHVCESEKAYKSOIPLPHGTHAFHACPCATQOLVCEONCIK 420
413 LIFGSPID 420
413 LIFGSPID 420
421 LIFGSPID 428
421 LIFGSPID 428

RESULT 6

PEL11 HUMAN
ID PEL11 HUMAN STANDARD; PRT; 418 AA.
AC Q96FA3; Q96SM0; Q96ZT5; Q9HCA0;
DT 29-MAR-2004 (Rel. 43, Created)
DT 29-MAR-2004 (Rel. 43, Last sequence update)
DT 10-MAY-2005 (Rel. 47, Last annotation update)
DE Pellino protein homolog 1 (Pellino 1) (Pellino related intracellular
DE signalling molecule).
GN Name=PEL11; Synonyms=PRISM;
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RC TISSUE=Thymus;
RX MEDLINE=21015032; PubMed=1132151; DOI=10.1007/s002510000249;
RA Rich T., Allen R.L., Lucas A.-M., Trowsdale J.,
RT "Pellino-related sequences from Caenorhabditis elegans and Homo
RT sapiens.";
Immunogenetics 52:145-149(2000).
[2]
RN NUCLEOTIDE SEQUENCE.
RX MEDLINE=21203570; PubMed=11306823;
RA Resch K., Jockusch H., Schmitt-John T.,
RT "Assignment of homologous genes, Pell1 and Pell2/PEL12, for the
RT Pelle adaptor protein Pellino to mouse chromosomes 11 and 14 and human
RT chromosomes 2p3.3 and 14q21, respectively, by physical and radiation
RT hybrid mapping.";
Cytogenet. Cell Genet. 92:172-174(2001).
[3]
RN NUCLEOTIDE SEQUENCE.
RA Kennedy E.J., Moynagh P.N.,
RT "PRISM, a novel mediator of Toll/IL-1 signalling.";
Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
[4]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].
RC TISSUE=Ovary, Placenta, and Testis;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heien F.,
RA Diatchenko L., Marusik K., Farmer A.A., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Donald M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Udén T.B., Toshitsuki S., Carninci P., Prange C.,
RA Raha S.S., Loguigliano N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Morley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
RA Fahney J., Helton E., Kettelman M., Madan A., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
RA Buttefield Y.S.N., Krzywinski M.I., Skalska U., Smalins D.E.,
RA Schmech A., Schein J.E., Jones S.J.M., Marra M.A.,
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.";
Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
[5]
RN NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA] OF 78-418.
RC TISSUE=Testis; carcinoma;
RX PubMed=14702039; DOI=10.1038/ng1285;
RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
RA Wakamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
RA Sekine M., Ohyashiki M., Nishi T., Shihahara T., Tanaka T., Ishii S.,
RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
RA Nagahara K., Murakami K., Yasuda T., Iwayanagi T., Wagatsuma M.,
RA Shiratori A., Sudo H., Hosoiri T., Kaku Y., Kodaira H., Kondo H.,
RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikawa E.,
RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa M.,

RA Yamazaki M., Minomiya K., Ishibashi T., Yamashita H., Murakawa K.,
RA Fujimori K., Tanai H., Kimata M., Watanabe M., Hiraoka S., Chiba Y.,
RA Ishida S., Ono Y., Takiguchi S., Watanabe S., Yoshida M., Hotta T.,
RA Kusano J., Kanehori K., Takahashi-Fuji A., Hara H., Tanase T.-O.,
RA Nomura Y., Togiyasu S., Komai F., Hara R., Takeuchi K., Arita M.,
RA Imose N., Kusashino K., Yuki H., Oshima A., Sasaki N., Aotsuka S.,
RA Yoshikawa Y., Matsunawa H., Ichihara T., Shiohata N., Sano S.,
RA Moriya S., Momiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
RA Nakagawa S., Senoh A., Mizoguchi H., Goto Y., Shimizu F., Wakabe H.,
RA Hishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
RA Fujimori Y., Komiyama M., Tashiro H., Tanigami A., Fujisawa T.,
RA Ono T., Yamada K., Fujii Y., Ozaki K., Hiroo M., Ohmori Y.,
RA Kawabata A., Hikihi T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
RA Okitani R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senba T.,
RA Matsumura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamaashita R.,
RA Nakai K., Yada T., Nakamura Y., Ohara O., Isogai T., Sugano S.;
RT "Complete sequencing and characterization of 21,243 full-length human
RT cDNAs.";
RL Nat. Genet. 36:40-45(2004).
RN [6]
RN FUNCTION, AND INTERACTIONS WITH IRAK1, IRAK4 AND TRAF6.
RP MEDLINE=2538429; PubMed=12496252; DOI=10.1074/jbc.M21212200;
RX Jiang Z., Johnson H.J., Nie H., Qin J., Bird T.A., Li X.;
RT "Pellino 1 is required for interleukin-1 (IL-1)-mediated signaling
RT through its interaction with the IL-1 receptor-associated kinase 4
RT (IRAK)-IRAK-tumor necrosis factor receptor-associated factor 6
RT (TRAF6) complex.";
RT J. Biol. Chem. 278:10952-10956(2003).
RN [7]
RN INTERACTIONS WITH TRAF6 AND MAP3K7.
RX MEDLINE=22689054; PubMed=12804775; DOI=10.1016/S0014-5793(03)00533-7;
RA Jensen L.E., Whitehead A.S.;
RT "Pellino2 activates the mitogen activated protein kinase pathway.";
RL FEBS Lett. 545:199-202(2003).
CC -I- FUNCTION: Scaffold protein involved in the IL-1 signaling pathway
CC via its interaction with the complex containing IRAK kinases and
CC TRAF6. Required for NF-kappa-B activation and IL-8 gene expression
CC in response to IL-1.
CC -I- SUBUNIT: Found in a complex containing TRAF6, IRAK1 and IRAK4.
CC -I- INTERACTS WITH MAP3K7.
CC -I- SIMILARITY: Belongs to the pellino family.
CC
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
CC EMBL: AJ278859; CAC04320.1; -; mRNA.
CC EMBL: AF302505; AAG15393.1; -; mRNA.
CC EMBL: AF300987; AAG17451.1; -; mRNA.
CC EMBL: BC011419; AAH11419.1; ALT. INT. mRNA.
CC EMBL: BC050019; AAH50019.1; -; mRNA.
CC EMBL: BC063611; AAH63611.1; -; mRNA.
CC EMBL: AK027668; BAB55280.1; ALT. INT. mRNA.
CC IntAct: Q96FA3; -;
CC DR Ensembl: ENSG00000197329; Homo sapiens.
CC DR HGNC: HGNC:8827; PEL1.
CC DR InterPro: IPR006800; Pellino.
CC DR PANTHER: PTHR12098; Pellino; 1.
CC DR Pfam: Pf04710; Pellino; 1.
CC FT CONFLICT 11 11 S -> F (in Ref. 1 and 5).
CC FT CONFLICT 260 260 S -> P (in Ref. 5).
CC SQ SEQUENCE 418 AA; 46286 MW; 233318A45E7546F7 CRC64;
CC
CC Query Match 83.7%; Score 1917; DB 1; Length 418;
CC Best Local Similarity 81.4%; Pred. No. 2,1e-151;
CC Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;

QY 1 MESPQGEHCAPNKEPVKYGELVLTGNGALPNGDRGRKRSAPALYKRPKANGKSTVH 60
DB 1 MFSPDQENH--PSKAPVKYGGELVLVIGNGSLPNGDGRKRSFPALFKRPKANGKSTVH 58
QY 61 VISTPQASKAISCKGHSISYTLSRNQTIVVEYTHDKDMDMVOVGSTSPDIPVDTVI 120
DB 59 IACTPQAAAKISCKDHSISYTLSRQIVVEYTHDSNTDMQDVGSTSPDIPVDTVI 118
QY 121 SGGQNTDEAQITQSTSRFRACRIVCDRNEPYTARIPAAFGDSSKNIFLEGKAAKMNKPD 180
DB 119 PGGQNSDQVSQSTISRFRACRIGERNPFTARITAAFGDSSKNIFLEGKAAKMTISG 178
QY 181 HMDGLTTNGVLVMPRGFTGTSRQPCWREISVCGDVTTLRETRSAQGRKLVESNTVL 240
DB 179 QMDGLTTNGVLVMPRGFTGTSRQPCWREISVCGDVTTLRETRSAQGRKLVESNTVL 238
QY 241 QDGSLLDLCATLMTAAGLFTTPQKILVLRGINARQCPGLNTLAPPSINRKE 300
DB 239 QDGSLLDLCATLMTAAGLFTTPQKILVLRGINARQCPGLNTLAPPSINRKE 298
QY 301 VVEKQPMAYLSGCHVGYHNMGRDTEANERECPCRTGVPYVPLMGCEAGFYVDAG 360
DB 299 VDEKQPMAYLNGCHVGYHNMGRDTEANERECPCRTGVPYVPLMGCEAGFYVDAG 358
QY 361 PPTAFTPCGHVCSKSAKYWSQIPPLPGHTAFAACPCATQVGEQNCIKLIFQGPID 420
DB 359 PPTAFTPCGHVCSKSAKYWSQIPPLPGHTAFAACPCATQVGEQNCIKLIFQGPID 418

RESULT 7
ID Q53T26 HUMAN PRELIMINARY; PRT; 418 AA.
AC Q53T26;
DT 13-SEP-2005 (T-EMBLrel. 31, Created)
DT 13-SEP-2005 (T-EMBLrel. 31, Last sequence update)
DE 13-SEP-2005 (T-EMBLrel. 31, Last annotation update)
GN Hypothetical protein PEL1.
OS Homo sapiens (human).
OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Hominidae;
OC Homo.
OX NCBI_TaxID=9606;
RN [1]
RN NUCLEOTIDE SEQUENCE.
RP Lesley K., Haekenson W., Martinka S.;
RT "The sequence of Homo sapiens BAC clone RP11-547M24.";
RL Submitted (OCT-1999) to the EMBL/GenBank/DBJ databases.
RN [2]
RN NUCLEOTIDE SEQUENCE.
RP Waterston R.H.;
RA Submitted (MAY-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RN NUCLEOTIDE SEQUENCE.
RA Waterston R.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [4]
RN NUCLEOTIDE SEQUENCE.
RP Wilson R.K.;
RL Submitted (APR-2005) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC012368; AAY24274.1; -; Genomic DNA.
KM Hypothetical protein.
SQ SEQUENCE 418 AA; 46286 MW; 233318A45E7546F7 CRC64;
CC
CC Query Match 83.7%; Score 1917; DB 2; Length 418;
CC Best Local Similarity 81.4%; Pred. No. 2,1e-151;
CC Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;
QY 1 MESPQGEHCAPNKEPVKYGELVLTGNGALPNGDRGRKRSAPALYKRPKANGKSTVH 60
DB 1 MFSPDQENH--PSKAPVKYGGELVLVIGNGSLPNGDGRKRSFPALFKRPKANGKSTVH 58
QY 61 VISTPQASKAISCKGHSISYTLSRNQTIVVEYTHDKDMDMVOVGSTSPDIPVDTVI 120

[illegible]

```

RA Shlratli, Maki K., Kawai J., Aizawa K., Arakawa T., Fukuda S.,
RA Hara A., Hashizume W., Imotani K., Ishii Y., Itoh M., Kagawa I.,
RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shimagawa A.,
RA Yaumishi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,
RA Bittney E., Hayashizaki Y.,
RT "Analysis of the mouse transcriptome based on functional annotation of
RT 60,770 full-length cDNAs.",
RL Nature 420:563-573(2002) .
RN [3]
RP NCBI/EBI/GENE/SEQUENCE [LARGE SCALE MRNA] .
RC TISSUE=Mammary tumor;
RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
RA Straubeberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
RA Klausner R.D., Collins K., Wagner L., Sherman C.M., Schuler G.D.,
RA Altschul S.F., Zeeberg B., Buettow K.H., Schaefer C.F., Bhat N.K.,
RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Hsieh P.,
RA Ditchenko L., Marsina K., Farmer A.C., Rubin G.M., Hong L.,
RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
RA Brownstein M.J., Uedini T.B., Tothlyki S., Carninci P., Prange C.,
RA Rana S.S., Loguella N.A., Peters G.J., Abramson R.D., Mullaly S.J.,
RA Bosak S.A., McGowan P.C., McKernan K.J., Malek J.A., Gunaratne P.H.,
RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Huylk S.W.,
RA Villalón D.K., Muzny D.M., Sodergren B.J., Lu X., Gibbs R.A.,
RA Bailey J., Helton E., Kettelman M., Mullan A.C., Rodriguez S., Sanchez A.,
RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
RA Rodriguez A.C., Gilmwood J., Schmutz J., Myers R.M.,
RA Butcherfield J.S.N., Krzywinski M.I., Skalecki M., Smalley D.E.,
RA Scherf A., Schein J.B., Jones S.J.M., Marra M.A.;
RT "Generation and initial analysis of more than 15,000 full-length human
RT and mouse cDNA sequences.",
RL Proc. Natl. Acad. Sci. U.S.A. 99:16699-16903(2002) .
CC -1- FUNCTION: Scaffold protein involved in the IL-1 signaling pathway
CC via its interaction with the complex containing IRAK kinases and
CC TRAF6. Required for NF-kappa-B activation and IL-8 gene expression
CC in response to IL-1 (By similarity).
CC -1- SUBUNIT: Found in a complex containing TRAF6, IRAK1 and IRAK4.
CC Interacts with MAP3K7 (By similarity).
CC -1- SIMILARITY: Belongs to the pellino family.
CC -----
CC This Swiss-Prot entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation --
CC the European Bioinformatics Institute. There are no restrictions on its
CC use as long as its content is in no way modified and this statement is not
CC removed.
CC -----
DR EMBL, AF302503; AAG5391.1; -; mRNA.
DR EMBL, AK009945; BAE2600.1; -; mRNA.
DR EMBL, AK045673; BAC32452.1; -; mRNA.
DR EMBL, AK076458; BAC36351.1; ALT INIT; mRNA.
DR EMBL, BC016515; AAH16515.1; -; mRNA.
DR Ensemble; ENSMUSG0000020134; Mus musculus.
DR MGI; MGI:191495; Pellino.
DR InterPro; IPR006800; Pellino.
DR PANTHER; PTHR12098; Pellino; 1.
DR Pfam; PF04710; Pellino; 1.
DR CONFLICT 5 7 DOB -> GSR (in Ref. 1).
SQ SEQUENCE 418 AA; 46259 MW; 26072654577BBP7 CRC64;
Query Match 83.5%; Score 1912; DB 1; Length 418;
Best Local Similarity 81.2%; Pred. No. 5.6e-151;
Matches 341; Conservative 39; Mismatches 38; Indels 2; Gaps 1;
OY 1 MFSPGSEHCAKPKPEPKYGEIVLVGNGALPLNGDGRRRSRFPALTYRPRKANGKPSYH 60
Db 1 MFSPDQGNH--PSAPPKYVELLVGLNGSLPLNGDRRRRSRFPALFRPRANGKYSTVH 58
OY 61 VISTPQASKAISKGGKSHSYTLTSLRNTVVVVEYTHDKDTLPMFOVGSTESPIDFVDTI 120
Db 59 IACIPQAKAKIISNDDGSHSYTLTSLRNTVVVVEYTHDKDTLPMFOVGSTESPIDFVDTV 118
OY 121 GSGQNDDEAGITOSTTIRFACRIVCDNEBYTARIFPAAGDSSKNIFLGKAAKMKPDG 180

```

Db 119 PGSQNSDTSQVOSTISRACRIICERSPPFTARIYAAGDSKNTFLGKAKAMKTSIDG 178
 QY 181 HMDGLTTNGVLVNHPRGFTBESOPGVWREISVCGDYVTLRETRSAQORGLVESETNVL 240
 Db 179 QMDGLTTNGVLVNHPRGFTBESOPGVWREISVCGDYVTLRETRSAQORGLVESETNVL 238
 QY 241 QDGLSLIDLCGATLLMTADGLFTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 300
 Db 239 QDGLSLIDLCGATLLMTADGLFTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 298
 QY 301 VVEEKOPMAYLSCGHVGHYNNMGHRSSTENARECEPMCRVGVYVPLMLGCEAGFYVDAG 360
 Db 299 VVDEKOPMAYLSCGHVGHYNNMGHRSSTENARECEPMCRVGVYVPLMLGCEAGFYVDAG 358
 QY 361 PPTHAFPCGHVSEKSAKYMSQIPLPHGTHAFPAACFPATQOLVGEONCIKILIFQGPID 420
 Db 359 PPTHAFPCGHVSEKSAKYMSQIPLPHGTHAFPAACFPATQOLVGEONCIKILIFQGPID 418

RESULT 9
 Q5SRW7_MOUSE PRELIMINARY; PRT; 418 AA.
 AC 05SRW7_MOUSE
 DT 10-MAY-2005 (TREMBlrel. 30, Created)
 DT 10-MAY-2005 (TREMBlrel. 30, last sequence update)
 DT 10-MAY-2005 (TREMBlrel. 30, last annotation update)
 DE Pellino 1.
 GN Name=Pell1; ORFNames=RP23-93011.1-001;
 OS Mus musculus (Mouse).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;
 OC Muridae; Murinae; Mus.
 NCBI_TaxId=10090;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RA Lovell J.;
 RL Submitted (FEB-2005) to the EMBL/GenBank/DBJ databases.
 DR EMBL; AL669979; CA126089.1; -; Genomic DNA.
 SQ SEQUENCE 418 AA; 46259 MW; 26072654577BBF7 CRC64;

Query Match 83.5%; Score 1912; DB 2; Length 418;
 Best Local Similarity 81.2%; Pred. No. 5.6e-151;
 Matches 341; Conservative 39; Mismatches 38; Indels 2; Gaps 1;

QY 1 MFSPOGEBCAPKEPVKYGELVVLGYNALPBGDGRKRSRPALYKRPKANGVKSTYH 60
 Db 1 MFSPODBNH--PSKAPVKYGEVLVLGYNGLPBGDGRKRSRPALYKRPKANGVKSTYH 58
 QY 61 VISTPOASKAISCKGHSISYTLISRNQTVVVEYTHDKOTDMPQVGSSTESPIDFVVTDTI 120
 Db 59 IACTPOAAKXISNKDHSISYTLISRAQTVVVEYTHDSNTDMPQIGRSTESPIDFVVTDTY 118
 QY 121 SSGQNTDEAQITOSTISRACRIICERNPEYTAIRIPAAGDSKNTFLGKAKAMKTPDG 180
 Db 119 PGSQNSDTSQVOSTISRACRIICERSPPFTARIYAAGDSKNTFLGKAKAMKTSIDG 178
 QY 181 HMDGLTTNGVLVNHPRGFTBESOPGVWREISVCGDYVTLRETRSAQORGLVESETNVL 240
 Db 179 QMDGLTTNGVLVNHPRGFTBESOPGVWREISVCGDYVTLRETRSAQORGLVESETNVL 238
 QY 241 QDGLSLIDLCGATLLMTADGLFTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 300
 Db 239 QDGLSLIDLCGATLLMTADGLFTPTOKHIEALROEINARPOCPVGLNTLAFPSINRKE 298
 QY 301 VVEEKOPMAYLSCGHVGHYNNMGHRSSTENARECEPMCRVGVYVPLMLGCEAGFYVDAG 360
 Db 299 VVDEKOPMAYLSCGHVGHYNNMGHRSSTENARECEPMCRVGVYVPLMLGCEAGFYVDAG 358
 QY 361 PPTHAFPCGHVSEKSAKYMSQIPLPHGTHAFPAACFPATQOLVGEONCIKILIFQGPID 420
 Db 359 PPTHAFPCGHVSEKSAKYMSQIPLPHGTHAFPAACFPATQOLVGEONCIKILIFQGPID 418

RESULT 10
 Q7ZXU3_XENLA PRELIMINARY; PRT; 418 AA.
 AC Q7ZXU3_XENLA
 DT 01-JUN-2003 (TREMBlrel. 24, Created)
 DT 01-JUN-2003 (TREMBlrel. 24, last sequence update)
 DT 01-OCT-2003 (TREMBlrel. 25, last annotation update)
 DE Pell1-prov protein.
 OS Xenopus laevis (African clawed frog).
 OC Eukaryota; Metazoa; Chordata; Vertebrata; Euteleostomi;
 OC Amphibia; Batrachia; Anura; Mesobatrachia; Pipiloidea; Pipidae;
 OC Xenopodinae; Xenopus; Xenopus.
 NCBI_TaxId=8355;
 RN [1]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Helel F.,
 RA Diatchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Uebli T.B., Toshlyuk S., Carninci P., Prange C.,
 RA Raha S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mallamy S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulyk S.W.,
 RA Villalon D.K., Muzny K.C., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Fahy J., Helton E., Kettelman M., Madan A.C., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A.C., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield V.S.N., Krzywinski M.I., Skalka U., Smalins D.E.,
 RA Scherch A., Schein J.E., Jones S.J.M., Maira W.A.,
 RT "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences."
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903 (2002).
 RN [2]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RX MEDLINE=22341132; PubMed=12454917; DOI=10.1002/vdy.10174;
 RA Klein S.L., Strausberg R.L., Wagner L., Pontius J., Clifton S.W.,
 RA Richardson P.;
 RT "Genetic and genomic tools for Xenopus research: The NIH Xenopus
 RT Initiative."
 RL Dev. Dyn. 225:384-391 (2002).
 RN [3]
 RP NUCLEOTIDE SEQUENCE.
 RC TISSUE=Embryo;
 RA Klein S., Strausberg R.;
 RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
 DR EMBL; BC044117; AAH44117.1; -; mRNA.
 DR InterPro; IPR006800; Pellino.
 DR Pfam; PF04710; Pellino; 1.
 SQ SEQUENCE 418 AA; 46231 MW; 04DCF12560BAD530 CRC64;

Query Match 82.3%; Score 1884; DB 2; Length 418;
 Best Local Similarity 80.5%; Pred. No. 1.2e-148;
 Matches 338; Conservative 39; Mismatches 41; Indels 2; Gaps 1;

QY 1 MFSPOGEBCAPKEPVKYGELVVLGYNALPBGDGRKRSRPALYKRPKANGVKSTYH 60
 Db 1 MFSPODBNH--PSKAPVKYGEVLVLGYNGLPBGDGRKRSRPALYKRPKANGVKSTYH 58
 QY 61 VISTPOASKAISCKGHSISYTLISRNQTVVVEYTHDKOTDMPQVGSSTESPIDFVVTDTI 120
 Db 59 IACTPOAAKXISNKDHSISYTLISRAQTVVVEYTHDSNTDMPQIGRSTESPIDFVVTDTY 118
 QY 121 SSGQNTDEAQITOSTISRACRIICERNPEYTAIRIPAAGDSKNTFLGKAKAMKTPDG 180
 Db 119 PGSQNSDTSQVOSTISRACRIICERNPEYTAIRIPAAGDSKNTFLGKAKAMKTLDD 178
 QY 181 HMDGLTTNGVLVNHPRGFTBESOPGVWREISVCGDYVTLRETRSAQORGLVESETNVL 240


```

Db 179 QMGGLTNGVLVHPRNGFEDSKGVWREISVCGNFSLETRRSAGQRKMWENESNEL 238
Qy 241 QDGSLLDLCATILMTADGLFHTPTIOKRIEALROEINARPOCVGLNTLAPPSINRKE 300
Db 239 QDGSLLDLCATILMTADGLFHTPTIOKRIEALROEINARPOCVGLNTLAPPSINRKE 298
Qy 301 VVEKOPMAVLSGCHVGHYNNMGHRSPTLANERECPCRTVGPVPLMTGCEAGFYVDAG 360
Db 299 VVDKOPWVYLNCGHVGYNHNGKBERDGDKEPCRCRVGPVPLMTGCEAGFYVDAG 358
Qy 361 PPTHAFPCGHCVCSEKSAKYWSOIPPLPHGTHAFHAACPFCATOLVGBQNCIKLIFQSPID 420
Db 359 PPTHAFPCGHCVCSEKSAKYWSOIPPLPHGTHAFHAACPFCATOLVGBQNCIKLIFQSPID 418

RESULT 11
052KX77 CHICK PRELIMINARY; PRT; 405 AA.
AC 052KX77
DT 25-OCT-2004 (TEMBLrel. 28, Created)
DT 25-OCT-2004 (TEMBLrel. 28, Last sequence update)
DE Hypochemical protein.
ORFNames=RCJMB04.944;
OS Gallus gallus (Chicken).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Archosauria; Aves; Neognathae; Galliformes; Phasianidae; Phasianinae;
OC Gallus.
OX NCBI_TaxID=9031;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RC STRAIN=CG; TISSUE=Bursa;
RA Caldwell R.B., Kierzek A.M., Arakawa H., Bezzubov Y., Zaim J.,
RA Fiedler P., Kuter S., Blagodatetski A., Kostovska D., Koter M.,
RA Plachy J., Carninci P., Hayashizaki Y., Buerstede J.M.;
RT "Full-length cDNAs from chicken bursa lymphocytes to facilitate
RT gene function analysis."
RL Genome Biol. 6:R6-R6(2005).
DR EMBL; AJ171997; CAG31656.1; -; mRNA.
DR InterPro; IPR006800; PelliNo.
DR Pfam; PF04710; PelliNo. 1.
KW Hypochemical protein.
SQ SEQUENCE 405 AA; 44790 MW; 8928PSESE52DB3BF CRC64;

Query Match 81.4%; Score 1864; DB 2; Length 405;
Best Local Similarity 81.8%; Pred. No. 5.5e-147;
Matches 333; Conservative 36; Mismatches 36; Indels 2; Gaps 1;

Qy 1 MSPGQGEHCAPNKEPVYKGLVYLGYNGALPNDGRGRKRSFPALYKRPKANGVPESTVH 60
Db 1 MSPGQGEHCAPNKEPVYKGLVYLGYNGALPNDGRGRKRSFPALYKRPKANGVPESTVH 58
Qy 61 VISTPQASKISCKGSHISYTLSSRNOTVVEYTHDQTMDFQVGRSTESPIDPVYTDI 120
Db 59 IACTPQAKAISNDQSHISYTLSSRAQTVVEYTHDNTMDFQGRSTESPIDPVYTDI 118
Qy 121 SSGSNTDEAOITOSTISRFACRIYCDNEPYTAIFAGFDSKNIFLGEKAAKMPDG 180
Db 119 PGSQNSNDQSVOSTISRFACRIYCDNEPYTAIFAGFDSKNIFLGEKAAKMPDG 178
Qy 181 HMDGLTNGVLVHPRNGFTEESOPGVWREISVCGDVYTLRETRSAQGRKLVESSTNL 240
Db 179 QMGGLTNGVLVHPRNGFTEESOPGVWREISVCGDVYTLRETRSAQGRKLVESSTNL 238
Qy 241 QDGSLLDLCATILMTADGLFHTPTIOKRIEALROEINARPOCVGLNTLAPPSINRKE 300
Db 239 QDGSLLDLCATILMTADGLFHTPTIOKRIEALROEINARPOCVGLNTLAPPSINRKE 298
Qy 301 VVEKOPMAVLSGCHVGHYNNMGHRSPTLANERECPCRTVGPVPLMTGCEAGFYVDAG 360
Db 299 VVDKOPWVYLNCGHVGYNHNGKBERDGDKEPCRCRVGPVPLMTGCEAGFYVDAG 358

```

```

Qy 361 PPTHAFPCGHCVCSEKSAKYWSOIPPLPHGTHAFHAACPFCATOLVGB 407
Db 359 PPTHAFPCGHCVCSEKSAKYWSOIPPLPHGTHAFHAACPFCATOLVGB 405

RESULT 12
04SGC9 TETNG PRELIMINARY; PRT; 428 AA.
AC 04SGC9
DT 13-SEP-2005 (TEMBLrel. 31, Created)
DT 13-SEP-2005 (TEMBLrel. 31, Last sequence update)
DE Chromosome 17 SCARF4597, whole genome shotgun sequence.
ORFNames=GSTENG00018701001;
OS Tetradon nigrovittatus (Green puffer).
OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;
OC Acanthomorphia; Acanthopterygii; Percormorphia; Tetraodontiformes;
OC Tetraodontidae; Tetraodontidae; Tetraodon.
OX NCBI_TaxID=99883;
RN [1]
RP NUCLEOTIDE SEQUENCE.
RA Jallion O., Aury J.M., Brunet F., Petit J.L., Stange-Thomann N.,
RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,
RA Micard S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,
RA Desliva C., Salanoubat M., Levy M., Boudet N., Castellano S.,
RA Auhouard V., Ublin C., Castelli V., Katinka M., Vacherie B.,
RA Blemont C., Skali Z., Cattolico L., Poulin J., De Bernardis V.,
RA Cnaud C., Duprat S., Broctier P., Coutanceau J.P., Gouzy J.,
RA Parra G., Lardier G., Chapple C., McKernan K.J., McEwan P., Bosak S.,
RA Kellie M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,
RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,
RA Laudet F., Schachter V., Quetier F., Saurin W., Scarpelli C.,
RA Wincker P., Lander E.S., Weissbach J., Roest Crollius H.;
RT "Genome duplication in the teleost fish Tetraodon nigrovittatus reveals
RT the early vertebrate proco-karyotype."
RL Nature 431:946-957(2004).
RN [2]
RP NUCLEOTIDE SEQUENCE.
RC GenomeScope; Whitehead Institute Centre for Genome Research;
RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.
CC -1- CAUTION: The sequence shown here is derived from an
CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is
CC preliminary data.
DR EMBL; CAAB01014597; CAG00303.1; -; Genomic DNA.
SQ SEQUENCE 428 AA; 47639 MW; E4187096D4454577 CRC64;

Query Match 78.1%; Score 1788; DB 2; Length 428;
Best Local Similarity 76.9%; Pred. No. 1.3e-140;
Matches 329; Conservative 38; Mismatches 53; Indels 8; Gaps 3;

Qy 1 MSPGQGEHCAPNKEPVYKGLVYLGYNGALPNDGRGRKRSFPALYKRPKANGVPESTVH 58
Db 1 MSPGQGEHCAPNKEPVYKGLVYLGYNGALPNDGRGRKRSFPALYKRPKANGVPESTVH 56
Qy 61 VISTPQASKISCKGSHISYTLSSRNOTVVEYTHDQTMDFQVGRSTESPIDPVYTDI 118
Db 59 IACTPQAKAISNDQSHISYTLSSRAQTVVEYTHDNTMDFQGRSTESPIDPVYTDI 116
Qy 121 SSGSNTDEAOITOSTISRFACRIYCDNEPYTAIFAGFDSKNIFLGEKAAKMPDG 180
Db 119 PGSQNSNDQSVOSTISRFACRIYCDNEPYTAIFAGFDSKNIFLGEKAAKMPDG 178
Qy 181 HMDGLTNGVLVHPRNGFTEESOPGVWREISVCGDVYTLRETRSAQGRKLVESSTNL 240
Db 179 QMGGLTNGVLVHPRNGFTEESOPGVWREISVCGDVYTLRETRSAQGRKLVESSTNL 238
Qy 241 QDGSLLDLCATILMTADGLFHTPTIOKRIEALROEINARPOCVGLNTLAPPSINRKE 300
Db 239 QDGSLLDLCATILMTADGLFHTPTIOKRIEALROEINARPOCVGLNTLAPPSINRKE 298
Qy 301 VVEKOPMAVLSGCHVGHYNNMGHRSPTLANERECPCRTVGPVPLMTGCEAGFYVDAG 360
Db 299 VVDKOPWVYLNCGHVGYNHNGKBERDGDKEPCRCRVGPVPLMTGCEAGFYVDAG 358

```


DB 301 HRKDPEKQPMWYLLQGHVGHYNNMGNPREBERGREGDRBECMCRKTPYVPLWLGCE 360

QY 353 AGFYVADGPTTHAFTPCGHVCESEKSAKCYMSQIPLPHGTHAFHACPCATOLVGEONCIK 412

DB 361 AGFYVADGPTTHAFTPCGHVCESEKSAKCYMSQIPLPHGTHAFHACPCATOLVGEONCIK 420

QY 413 LIFOGPID 420

DB 421 LIFOGPID 428

RESULT 13

Q4SBZ1.TETNG PRELIMINARY; PRT; 440 AA.

ID Q4SBZ1.TETNG

AC Q4SBZ1.TETNG

DT 13-SEP-2005 (TEMBLrel. 31, Created)

DT 13-SEP-2005 (TEMBLrel. 31, Last sequence update)

DT 13-SEP-2005 (TEMBLrel. 31, Last annotation update)

DE Chromosome 14 SCAF14660, whole genome shotgun sequence.

DE (Fragment).

GN ORFNames=GSTENG0020745001;

OS Tetraodon nigroviridis (Green puffer).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Actinopterygii; Neopterygii; Teleostei; Euteleostei; Neoteleostei;

OC Acanthomorpha; Acanthopterygii; Percormorpha; Tetraodontiformes;

OC Tetraodontidae; Tetraodontidae; Tetraodon.

NCBI_TaxID=99883;

RA NUCLEOTIDE SEQUENCE.

RA Jallion O., Aury J.M., Brunet F., Petit J.L., Strange-Thomann N.,

RA Mauceli E., Bouneau L., Fischer C., Ozouf-Costaz C., Bernot A.,

RA Micard S., Jaffe D., Fisher S., Lutfalla G., Dossat C., Segurens B.,

RA Dasilva C., Salanoubat M., Levy M., Boudet N., Castellano S.,

RA Anthouard V., Jubin C., Castell J.V., Karinka M., Vacherie B.,

RA Blemont C., Skalli Z., Cattolico L., Poulain J., De Berardinis V.,

RA Crenaud C., Duprat S., Brottier P., Couanceau J.P., Gouzy J.,

RA Parra G., Lardier G., Chapelle C., McKernan K.J., McEwan P., Bosak S.,

RA Kellis M., Wolff J.N., Guigo R., Zody M.C., Mesirov J.,

RA Lindblad-Toh K., Birren B., Nusbaum C., Kahn D., Robinson-Rechavi M.,

RA Lander P., Schachter V., Quetier F., Saurin W., Scapellato C.,

RA Wincker P., Lander E.S., Weissenbach J., Roest Crolius H.,

RT Genome duplication in the teleost fish Tetraodon nigroviridis reveals

RL the early vertebrate proto-karyotype.";

RL Nature 431:946-957(2004).

RA [2]

RA NUCLEOTIDE SEQUENCE.

RA Genoscope, Whitehead Institute Centre for Genome Research;

RL Submitted (FEB-2004) to the EMBL/GenBank/DBJ databases.

CC -!- CAUTION: The sequence shown here is derived from an

CC EMBL/GenBank/DBJ whole genome shotgun (WGS) entry which is

CC preliminary data.

DR EMBL; CA01014660; CAC01841.1; -; Genomic_DNA.

FT NON_TER 1

SQ SEQUENCE 440 AA; 48445 MW; A2DB6C7CF9E85099 CRC64;

Query Match 71.9%; Score 1647; DB 2; Length 440;

Best Local Similarity 67.9%; Pred. No. 8.4e-129; Indels 32; Gaps 4;

Matches 299; Conservative 47; Mismatches 62;

QY 12 PNKEPVKGLVVLGNGALPNDGRGRKSRPALYKPKANGVPSFVHISTFQASKAI 71

DB 4 PVMDPIKGLVILGNGSLPSGDRGRKSRPALYKPKANGVPSFVHILNSPQDKAV 63

QY 72 SCGQHSISTLRNQNQVVEYTHDKQTMFQVGRSTESISIDFVYDTTIGSQNTDEAOI 131

DB 64 HSSGQHSISFTLRNQNQVVEYCHDNTDMFQIGRSTESPIDFVYDTTIGSAGXDGPSTI 123

QY 132 TOSTISFACRIYCDREPEYTAIFAAGFSSKNIIFGEAKAKWPDGADGLTNGVL 191

DB 124 APSTISFACRIYCDREPEYTAIFAAGFSSKNIIFGEAKAKWPDGADGLTNGVL 183

QY 192 VMHPRGFTESQPGVWRISVCGDVYTLRETRSAOQRKLV----- 233

DB 184 VMHPRGFTESQPGVWRISVCGDVYTLRETRSAOQRKLV----- 242

QY 234 ---ESEFNVLQDGLSLDLCGATLMTADGLFTTPQKHIEALROETINARPCPVGLNT 290

DB 243 LQAEGETSALRDSLDVLCGATLMTADGLFTTPQKHIEALROETINARPCPVGLNT 302

QY 291 LAEPISNRKEVNEKQPMWYLLQGHVGHYNNMGNPREBERGREGDRBECMCRKTPYVPLWLGCE 360

DB 303 LAFPSLPRSHSLERQPMWYLLQGHVGHYNNMGNPREBERGREGDRBECMCRKTPYVPLWLGCE 362

QY 342 GPVYPLMLGCEGAFYVDAGPTTHAFTPCGHVCESEKSAKCYMSQIPLPHGTHAFHACPCPCA 401

DB 363 GPVYPLMLGCEGAFYVDAGPTTHAFTPCGHVCESEKSAKCYMSQIPLPHGTHAFHACPCPCA 422

QY 402 TOLVGEONCIKLFQGPID 420

DB 423 AAL-GSPGMTRLLFQGPID 440

RESULT 14

P6L13_MOUSE

ID P6L13_MOUSE STANDARD; PRT; 445 AA.

AC Q6BKX6; Q6K329;

DT 29-MAR-2004 (Rel. 43, Created)

DT 29-MAR-2004 (Rel. 43, Last sequence update)

DT 13-SEP-2005 (Rel. 48, Last annotation update)

DE Pellino protein homolog 3 (Pellino 3).

GN Name=Pellino3;

OS Mus musculus (Mouse).

OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;

OC Mammalia; Eutheria; Euarchontoglires; Glires; Rodentia; Sciurognathi;

OC Muridae; Muridae; Murinae; Mus.

NCBI_TaxID=10090;

RA [1]

RA NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RC STRAIN=C57BL/6J; TISSUE=Retina;

RX MEDLINE=22354683; PubMed=12466851; DOI=10.1038/nature01266;

RA Okazaki I., Furuno M., Kasukawa T., Adachi Y., Bono H., Kondo S.,

RA Nikaide I., Furuno M., Saito R., Suzuki H., Yamanaka I., Kiyosawa H.,

RA Yagi K., Tomaru Y., Hasegawa Y., Nogami A., Schonbach C., Gotohori T.,

RA Balderelli R., Hill D.P., Bult C., Hume D.A., Quackenbush J.,

RA Schirml L.M., Kanapin A., Matsuda H., Batelov S., Beisel K.W.,

RA Blake J.A., Bird T., Brusic V., Choithia C., Corbani L.E., Cousins S.,

RA Dalla B., Dragani T.A., Fletcher C.F., Forrest A., Fraser K.S.,

RA Gaasterland T., Gariboldi M., Gissi C., Godzik A., Gough J.,

RA Grimmond S., Guenichich S., Hirokawa N., Jackson I.J., Jarvis E.D.,

RA Kanai A., Kawaji H., Kawasawa Y., Kedzierski R.M., King B.L.,

RA Konagaya A., Kurochkin I.V., Lee Y., Lenhard B., Lyons P.A.,

RA Maglocot D.R., Maltais L., Marchionni L., McKenzie L., Miki H.,

RA Nagashima T., Numata K., Okido T., Pavan W.D., Pereira G., Pebole G.,

RA Petrovsky N., Pillai R., Pontius J.U., Qi D., Ramachandran S.,

RA Ravasi T., Reed J.C., Reed D.J., Reid J., Ring B.Z., Ringwald M.,

RA Sandelin A., Schneider C., Sempke C.A., Seton M., Shimada K.,

RA Sultana R., Takenaka Y., Taylor M.S., Teasdale R.D., Tomita K.,

RA Verardo R., Wagner L., Wahlestedt C., Wang Y., Watanabe Y., Wells C.,

RA Wilming L.G., Wymshaw-Boris A., Yanagisawa M., Yang I., Yang L.,

RA Yuan Z., Zaytoun M., Zhu Y., Zimmer A., Carninci P., Hayatsu N.,

RA Hirose K., Kishikawa T., Kono H., Nakamura M., Sakazume N., Sato K.,

RA Shiraki T., Waki K., Kawai J., Aizawa K., Arikawa T., Fukuda S.,

RA Hara A., Haseizume W., Imotani K., Ishii Y., Itoh K., Kagawa I.,

RA Miyazaki A., Sakai K., Sasaki D., Shibata K., Shingawa A.,

RA Yabuuchi A., Yoshino M., Waterston R., Lander E.S., Rogers J.,

RA Birney E., Hayashizaki Y.,

RT "Analysis of the mouse transcriptome based on functional annotation of

RT 60,770 full-length cDNAs."

RL Nature 420:563-573(2002).

RA [2]

RA NUCLEOTIDE SEQUENCE [LARGE SCALE MRNA].

RC TISSUE=Mammary tumor;

RX MEDLINE=22368257; PubMed=12477932; DOI=10.1073/pnas.242603899;

RA Strausberg R.L., Feingold E.A., Grouse L.H., Derge J.G.,

RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,

RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,

RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,
 RA Stapleton M., Soares M.B., Bonaldo M.F., Casavant T.L., Scheetz T.E.,
 RA Brownstein M.J., Udell T.B., Toshiyuki S., Carninci P., Prange C.,
 RA Rana S.S., Loquellano N.A., Peters G.J., Abramson R.D., Mullany S.J.,
 RA Bosak S.A., McEwan P.J., McKernan K.J., Malek J.A., Gunaratne P.H.,
 RA Richards S., Worley K.C., Hale S., Garcia A.M., Gay L.J., Hulys S.W.,
 RA Villalón D.R., Muzny D.M., Sodergren E.J., Lu X., Gibbs R.A.,
 RA Feby J., Helton E., Kettman M., Madan A., Rodriguez S., Sanchez A.,
 RA Whiting M., Madan A., Young A.C., Shevchenko Y., Bouffard G.G.,
 RA Blakesley R.W., Touchman J.W., Green E.D., Dickson M.C.,
 RA Rodriguez A.C., Grimwood J., Schmutz J., Myers R.M.,
 RA Butlerfield Y.S.N., Krzyzanski M.I., Skalska U., Smalilus D.E.,
 RA Schnerch A., Schein J.E., Jones S.J.M., Marra M.A.,
 RA "Generation and initial analysis of more than 15,000 full-length human
 RT and mouse cDNA sequences".
 RL Proc. Natl. Acad. Sci. U.S.A. 99:16899-16903(2002).
 CC -1- FUNCTION: Scaffold protein which probably links Toll-like
 CC receptors (TLRs) to basic cellular processes via its interaction
 CC with the complex containing IRAK kinases and TRAF6. Can activate
 CC AP1/JUN and ELK1. Not required for NF-kappa-B activation. Probably
 CC involved in innate immune response (By similarity).
 CC -1- SUBUNIT: Interacts with TRAF6, IRAK1, MAP3K14 and MAP3K7 (By
 CC similarity).
 CC -1- SIMILARITY: Belongs to the pellino family.

 CC This Swiss-Prot entry is copyright. It is produced through a collaboration
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation
 CC the European Bioinformatics Institute. There are no restrictions on its
 CC use as long as its content is in no way modified and this statement is not
 CC removed.
 CC
 CC EMBL: AK044418; BAC13113.1; -; mRNA.
 CC EMBL: BC028931; AAH28931.1; -; mRNA.
 CC Ensemble: ENSMUSG0000024901; Mus musculus.
 CC MGI: MGI:1924963; 6030441P14Rik.
 CC InterPro: IPR006800; Pellino; 1.
 CC PANTHER: PTHR12098; Pellino; 1.
 CC Pfam: PF04710; Pellino; 1.
 CC CONFLICT 353 353 R -> Q (in Ref. 1).
 CC FT SEQUENCE 445 AA; 48168 MW; 4C09BE147D5BD941 CRC64;
 SQ
 Query Match 71.3%; Score 1632; DB 1; Length 445;
 Match Local Similarity 71.1%; Pred. No. 1.5e-127;
 Matches 297; Conservative 44; Mismatches 73; Indels 4; Gaps 3;
 QY 4 PGOEHCAPKPEPVYKGLVYVYNGALPNDGRGRKSRPALYKPRKANGKYSTVAVIS 63
 DB 31 PGBE--ALAGEBPKYGEELVYVYNGALSGDKRRSRRLALSRPRANGKYPVYVHIS 88
 QY 64 TPQASKAISCKGQHSISYTLISRNQTVVVEYTHDKDTMFOVGRSTESPIDFVVTITSGS 123
 DB 89 TPLVSKALSRGQHSISFTLSRSHSVIVETVTHDSKDMFOIGRSTEMIDFVVTITSGS 148
 QY 124 QNTDEAQTITSTISRFRCIVCDRENEYTARIPAGDSSKNITFLGKAAKAKNPDGMD 183
 DB 149 GAT-EGPSAOSTISRYACRICIDRRPYTARIYAAGDASNIFLGRRAKWRPDDIAMD 207
 QY 184 GLTTNGVLVNHPRGPGFPEESOPGWREISYCGDYVTLRFRPSAQRKLVESSENVVADG 243
 DB 208 GLTTNGVLVNHPRGPGFPEESOPGWREISYCGDYVTLRFRPSAQRKLVESSENVVADG 267
 QY 244 SLIDLCGATLLMRTADGLFTPTQKHEALROEINARPOCPVGLNTIAPPSINRKEVE 303
 DB 268 SLIDLCGATLLMRTADGLFTPTQKHEALROEINARPOCPVGLNTIAPPSINRKEVE 327
 QY 304 EK-QPMAYLSCGHVGHNGHSDTEANRECEMCTGVPPYPLMGTCEAGFVVDGPP 362
 DB 328 DKQOPWYVYVCGHGHGCHGCRERERPOREBCLCLVPEYVPLMGTCEAGFVVDGPP 387
 QY 363 THAFTPCGHVCEKSAKYMSOIPLEPHGTHAFHACPCFCAQVLEKNCIKILFOGSPID 420
 DB 388 SHAFAPCGHVCSEKTIARYMOTPLPHGTHAFHACPCFCAQVLEKNCIKILFOGSPID 445

RESULT 15
 ID PEL13_HUMAN STANDARD; PRT; 469 AA.
 AC Q8N2H9; Q8N3E1; Q8N9Q6; Q8TAM7; Q8TED5;
 DT 29-MAR-2004 (Rel. 43, Created)
 DT 29-MAR-2004 (Rel. 43, Last sequence update)
 DT 13-SEP-2005 (Rel. 48, Last annotation update)
 DE Pellino protein homolog 3 (Pellino 3).
 GN Name=PEL13;
 OS Homo sapiens (Human).
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
 OC Mammalia; Eutheria; Euarchontoglires; Primates; Catarrhini; Homidae;
 OC Homo.
 OX NCBI_TaxID=9606;
 RN [1]
 RP NUCLEOTIDE SEQUENCE (ISOFORMS 1 AND 2), FUNCTION, TISSUE SPECIFICITY,
 RP AND INTERACTIONS WITH IRAK1, TRAF6, MAP3K7 AND MAP3K14.
 RX MEDLINE=22756745; PubMed=12874243;
 RA Jensen L.E., Whitehead A.S.;
 RT "Pellino3, a novel member of the Pellino protein family, promotes
 RT activation of c-Jun and Elk-1 and may act as a scaffolding protein.";
 RL J. Immunol. 171:1500-1506(2003).
 RN [2]
 RP NUCLEOTIDE SEQUENCE (LARGE SCALE MRNA) (ISOFORM 1).
 RC TISSUE=Adipose;
 RG The German cDNA consortium;
 RL Submitted (JUL-2002) to the EMBL/GenBank/DBJ databases.
 RN [3]
 RP NUCLEOTIDE SEQUENCE (LARGE SCALE MRNA) (ISOFORMS 1 AND 3).
 RC TISSUE=Adipose tissue; Thyroid, and Uterus;
 RX PubMed=14702039; DOI=10.1038/ng1285;
 RA Ota T., Suzuki Y., Nishikawa T., Otsuki T., Sugiyama T., Irie R.,
 RA Makamatsu A., Hayashi K., Sato H., Nagai K., Kimura K., Makita H.,
 RA Sekite M., Oobayashi M., Nishi T., Shibahara T., Tanaka T., Ishii S.,
 RA Yamamoto J.-I., Saito K., Kawai Y., Isono Y., Nakamura Y.,
 RA Nagahara K., Murakami K., Yaeuda T., Iwayanagi T., Wagatsuma M.,
 RA Shiratori A., Sudo H., Hosobori T., Kaku Y., Kodaira H., Kondo H.,
 RA Sugawara M., Takahashi M., Kanda K., Yokoi T., Furuya T., Kikkawa E.,
 RA Omura Y., Abe K., Kamihara K., Katsuta N., Sato K., Tanikawa K.,
 RA Yamazaki M., Ninomiya K., Ishibashi T., Yamashita H., Murakawa K.,
 RA Fujimori K., Tanai H., Kimura M., Watanabe S., Yoshida M., Hattori Y.,
 RA Ishida S., Oono Y., Takiguchi S., Watanabe S., Yosida M., Hotuta T.,
 RA Kusano J., Kanehori K., Takahashi-Fujii A., Hara H., Tanase T.-O.,
 RA Nomura Y., Matsuhira K., Yuki H., Oshima A., Sasaki N., Arita M.,
 RA Nomura S., Wakiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Moriya S., Nomiyama H., Satoh N., Takami S., Terashima Y., Suzuki O.,
 RA Nishigaki H., Watanabe T., Sugiyama A., Takemoto M., Kawakami B.,
 RA Yamazaki M., Watanabe K., Kumagai A., Itakura S., Fukuzumi Y.,
 RA Fujimori Y., Komiyama M., Tashiro H., Taniguchi A., Fujitawa T.,
 RA Ono T., Yamada K., Fujii Y., Ozaki K., Hirao M., Ohmori Y.,
 RA Kawabata A., Hikiji T., Kobatake N., Inagaki H., Ikema Y., Okamoto S.,
 RA Ohtsuki R., Kawakami T., Noguchi S., Itoh T., Shigeta K., Senda T.,
 RA Matsunura K., Nakajima Y., Mizuno T., Morinaga M., Sasaki M.,
 RA Togashi T., Oyama M., Hata H., Watanabe M., Komatsu T.,
 RA Mizushima-Sugano J., Satoh T., Shirai Y., Takahashi Y., Nakagawa K.,
 RA Okumura K., Nagase T., Nomura N., Kikuchi H., Masuho Y., Yamashita R.,
 RA Nakai K., Yada T., Nakamura Y., Ohara O., Iwaga T., Sugano S.,
 RT "Complete sequencing and characterization of 21,243 full-length human
 RT cDNAs.";
 RL Nat. Genet. 36:40-45(2004).
 RN [4]
 RP NUCLEOTIDE SEQUENCE (LARGE SCALE MRNA) (ISOFORMS 2 AND 4).
 RC TISSUE=Brain, and Lung;
 RX MEDLINE=22388257; PubMed=12477932; DOI=10.1073/pnas.242603899;
 RA Strausberg R.L., Feingold E.A., Grouse L.H., Dege J.G.,
 RA Klausner R.D., Collins F.S., Wagner L., Shenmen C.M., Schuler G.D.,
 RA Altschul S.F., Zeeberg B., Buetow K.H., Schaefer C.F., Bhat N.K.,
 RA Hopkins R.F., Jordan H., Moore T., Max S.I., Wang J., Heieh F.,
 RA Datchenko L., Marusina K., Farmer A.A., Rubin G.M., Hong L.,

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioacceleration Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2006, 20:56:21 ; Search time 40 Seconds
(without alignments)
1010.276 Million cell updates/sec

Title: US-10-041-030-4
Perfect score: 2290
Sequence: 1 MRSFGQERHCAFKPKYK.....ATQVGEQNCIKLFGQPID 420

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 283416 seqs, 9621673 residues
Total number of hits satisfying chosen parameters: 283416

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :
1: p1r1:*
2: p1r2:*
3: p1r3:*
4: p1r4:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|-------|-------------|--------|----------|----------------------|
| 1 | 890 | 38.9 | 484 | 2 T25731 | hypothetical prote |
| 2 | 122.5 | 5.3 | 215 | 2 T28405 | ORF MSV244 hypothe |
| 3 | 103.5 | 4.5 | 1591 | 2 A54146 | invasion-inducing |
| 4 | 100 | 4.4 | 350 | 1 T03447 | dihydroxyacetophenol |
| 5 | 98 | 4.3 | 994 | 2 A10072 | probable autotrans |
| 6 | 95.5 | 4.2 | 978 | 2 C86547 | polymorphic outer |
| 7 | 95.5 | 4.2 | 978 | 2 B81593 | polymorphic membra |
| 8 | 95.5 | 4.2 | 978 | 2 G72076 | polymorphic outer |
| 9 | 92.5 | 4.0 | 874 | 2 C81804 | alanine-tRNA ligase |
| 10 | 92 | 4.0 | 3051 | 2 S42373 | hypothetical prote |
| 11 | 91.5 | 4.0 | 576 | 2 T40476 | hypothetical prote |
| 12 | 91.5 | 4.0 | 1449 | 2 A12017 | hypothetical prote |
| 13 | 90.5 | 4.0 | 520 | 2 B33830 | cation efflux syst |
| 14 | 90.5 | 4.0 | 823 | 2 F90914 | hypothetical prote |
| 15 | 90 | 3.9 | 449 | 1 NBHDS | complement factor |
| 16 | 90 | 3.9 | 1119 | 2 S41647 | zinc finger 5 prot |
| 17 | 90 | 3.9 | 1119 | 2 A88481 | protein C16A3.6 [1 |
| 18 | 90 | 3.9 | 1231 | 1 NBHDS | complement factor |
| 19 | 90 | 3.9 | 1996 | 2 F71405 | probable TWV resis |
| 20 | 89.5 | 3.9 | 823 | 2 G90848 | probable exonuclea |
| 21 | 89.5 | 3.9 | 823 | 2 E85706 | probable exonuclea |
| 22 | 89 | 3.9 | 288 | 2 A10942 | probable transmem |
| 23 | 89 | 3.9 | 681 | 2 AD3318 | DNA primase (EC 2. |
| 24 | 89 | 3.9 | 747 | 2 T36812 | probable dehydroge |
| 25 | 89 | 3.9 | 2098 | 2 T13166 | rough deal protein |
| 26 | 89 | 3.9 | 2265 | 1 FNBO | fibronectin - bovi |
| 27 | 89 | 3.9 | 2458 | 2 T17470 | probable polyketid |
| 28 | 89 | 3.9 | 2660 | 2 E85822 | probable invasin 2 |
| 29 | 88.5 | 3.9 | 558 | 2 T02704 | hypothetical prote |

| | | | | | |
|----|------|-----|------|----------|--------------------|
| 30 | 88.5 | 3.9 | 1104 | 2 I38869 | transcription fact |
| 31 | 88.5 | 3.9 | 1121 | 2 S54504 | hypothetical prote |
| 32 | 88 | 3.8 | 812 | 2 AC2349 | hypothetical prote |
| 33 | 88 | 3.8 | 1345 | 2 H90975 | hypothetical prote |
| 34 | 87.5 | 3.8 | 520 | 2 JC4699 | cadmium, zinc, cob |
| 35 | 87.5 | 3.8 | 724 | 2 A48569 | antigen Em100 - B1 |
| 36 | 87.5 | 3.8 | 1650 | 2 S28721 | hypothetical prote |
| 37 | 87 | 3.8 | 375 | 2 T03681 | putrescine N-methy |
| 38 | 87 | 3.8 | 923 | 2 G81253 | probable oxidoredu |
| 39 | 86.5 | 3.8 | 671 | 2 JE0288 | krueppel-type zinc |
| 40 | 86 | 3.8 | 518 | 1 NIBCBT | nitrogenase (EC 1. |
| 41 | 86 | 3.8 | 610 | 2 T22687 | hypothetical prote |
| 42 | 86 | 3.8 | 810 | 2 C84845 | probable salt-indu |
| 43 | 86 | 3.8 | 1326 | 2 B83770 | hypothetical prote |
| 44 | 86 | 3.8 | 1355 | 2 S46224 | peroxidase - frul |
| 45 | 86 | 3.8 | 1714 | 1 S18644 | multifunctional am |

ALIGNMENTS

| | | |
|-----------------------|--|--|
| RESULT 1 | T25731 | hypothetical protein F25B4.2 - Caenorhabditis elegans |
| C/Species: | Caenorhabditis elegans | |
| C/Date: | 15-Oct-1999 | #sequence_revision 15-Oct-1999 #ext_change 31-Dec-2004 |
| C/Accession: | T25731 | |
| R/Gattung: | S. | |
| A: | Submitted to the EMBL Data Library, July 1996 | |
| A: | Description: The sequence of C. elegans cosmid F25B4. | |
| A: | Reference number: Z28076 | |
| A: | Accession: T25731 | |
| A: | Status: preliminary; translated from GB/EMBL/DBJ | |
| A: | Molecule type: DNA | |
| A: | Residues: 1-484 <GAT> | |
| A: | Cross-references: UNIPROT:Q22967; UNIPARC:UPI0000179271; EMBL:U64842; P1DN:AA837081.1; | |
| A: | Experimental source: strain Bristol NZ; clone F25B4 | |
| C/Genetics: | | |
| A: | Gene: CESP:F25B4.2 | |
| A: | Map position: 5 | |
| A: | Introns: 35/2; 78/3; 112/3; 136/1; 191/3; 334/3; 406/2; 437/3 | |
| Query Match | 38.9%; Score 890; DB 2; Length 484; | |
| Best Local Similarity | 43.5%; Pred. No. 1.7e-64; | |
| Matches | 178; Conservative 68; Mismatches 143; Indels 20; Gaps 7; | |
| QY | 19 YGELVVLGYNGALPNDGRGR-KSRPALYKPRANGVKSTVYVISTPOASKAISKGGH 77 | |
| DB | 28 YGELILGFNGQAKENRATSKRYITKVKLRKRSANQIKCTYANVSTSD-TKLTOKARI 86 | |
| QY | 78 SISYTLRSNQVVEVYTHDKDTDFQVGRSTESPIDFVVVTDI-IGSSQNTDEN----QI 131 | |
| DB | 87 TVSFHSDSKMSVYIEAADPSKDMFQIGRASDDQIDFVYIDTWMFLPERSDAVPAAPQI 146 | |
| QY | 132 -----TGSTSRKRCIVCDRNEPYTARIPAGDSSKNITFLGKRAKKNPDHMG 184 | |
| DB | 147 DYLEKDRSTISRFACRILIDRNSNKAYLVAAAGDAQNISINKSKMTKSNBEVDG 206 | |
| QY | 185 LITNGVLLVNHPRG----GFTESSQPVWREISVCGVYTLRETRSAQGRKLVESSTNV 240 | |
| DB | 207 LITNGVLLVNHPRG----GFTESSQPVWREISVCGVYTLRETRSAQGRKLVESSTNV 266 | |
| QY | 241 QDGLSLDLCGATLMTADGLFTPTQKIEALROEINARPOCPVGLNTLAPPSINRKE 300 | |
| DB | 267 QDGLTLIDLCGATLMTADGLERSFGRELEVALDRINAGRCQPVNLTVLVPKRNGR 326 | |
| QY | 301 VAEKQPMAYLSGHHYGHYHNGHSDTEANRBCRCMCCTVGP-VYPLVLCGAGGYVDA 359 | |
| DB | 327 QINRRQPYVYLQCGHQGHGHEWGVQENSQGRSKCPICLVESBRITQLSMGEPSFHDLS 386 | |
| QY | 360 GPRTAFTPOCGHVCSEKSAKYSQPLPHGTAAFAHACFPACATQVGEQ 408 | |
| DB | 387 GVIDHTFPGHMAKQIVLWMSRIPLPGTGTCTRYDVEVCFCTQLATLR 435 | |

RESULT 2

T28405
ORF MSV244 hypothetical protein - Melanoplus sanguinipes entomopoxvirus
C/Species: Melanoplus sanguinipes entomopoxvirus
C/Date: 21-Jan-2000 #sequence_revision 21-Jan-2000 #text_change 09-Jul-2004
C/Accession: T28405
R/Alonso, C.L.; Tulman, E.R.; Lu, Z.; Oms, E.; Kuteh, G.F.; Rock, D.L.
J. Virol. 73, 533-552, 1999
A/Title: The genome of Melanoplus sanguinipes entomopoxvirus.
A/Reference number: Z20484; MUID:99102612; PMID:9847359
A/Accession: T28405
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-215 <ARO>
A/Cross-references: UNIPROT:Q9YVJ8; UNIPARC:UPI00000F81BF; EMBL:AF063866; NID:g4049647;
C/Genetic1:
A/Note: MSV244

Query Match 5.3%; Score 122.5; DB 2; Length 215;
Best Local Similarity 18.7%; Pred. No. 0.016;
Matches 61; Conservative 31; Mismatches 84; Indels 151; Gaps 13;

QY 81 YTLRNQTVVEYTHDKDTMFQVGRSTESPIDFVVTDTISGSQNTDEAOITOSTISRFA 140
DB 19 YNLIYNNK-----NYT-----FTIGRN--SKCDIFIND-----KKISNIA 50
QY 141 CRIVCDNRPETPTARIFPAQFDSSKNITLGEKAKAKNPDGMGLTTNGVLVNHPRGGFT 200
DB 51 CIHKCD-YEKKKCFIYGCSNKKYLYPDNN-----DGLQNGIFIKFPN----- 94
QY 201 EESQPGVWRREISVCGDYVTLRETSAAQQRGLVSESTNVLDGSLDLCGATLLMRTADG 260
DB 95 -----LEMDLSTINGNITYL-----KSNIKMLINELDTLLDISGNIYWRSI-- 139
QY 261 LFTPTOKHLEALROEINAAAPQCPVGLNTLAFPSINRKEVEBKOPWAYLSCGHVGYH 320
DB 140 ----- 139
QY 321 NWGHRSTTEANEREC--PKCRTPGVYPLWLCGAGFYVDAGPPTAFTPGHVCSEKSA 378
DB 140 -----DVKLNKKICNNSMC-TSTKY-----EYMFYKCGHKINKKTA 174
QY 379 KYWSQI-----PLPHGTHAFHAACDFC 400
DB 175 TFWKTKTMLCNDITYNKKLINTCPFC 201

RESULT 3

A54146
invasion-inducing protein Tiam-1 - mouse
C/Species: Mus musculus (house mouse)
C/Date: 02-Aug-1994 #sequence_revision 02-Aug-1994 #text_change 09-Jul-2004
C/Accession: A54146
R/Habers, G.G.M.; Scholtes, E.H.M.; Zuydgeest, D.; van der Kammen, R.A.; Stam, J.C.; Bet
Cell 77, 537-549, 1994
A/Title: Identification of an invasion-inducing gene, Tiam-1, that encodes a protein wit
A/Reference number: A54146; MUID:94243921; PMID:7999144
A/Accession: A54146
A/Status: preliminary
A/Molecule type: mRNA
A/Residues: 1-1591 <HAB>
A/Cross-references: UNIPROT:Q60610; UNIPARC:UPI0000028188; GB:U05245; NID:g497638; PIDN:
F;1040-1234/Domain: CDC24 homology <CD24>

Query Match 4.5%; Score 103.5; DB 2; Length 1591;
Best Local Similarity 22.1%; Pred. No. 8.7;
Matches 81; Conservative 53; Mismatches 166; Indels 67; Gaps 18;

QY 3 SPGEHECAPNKEVKYGEIVLVINGALPNGDRGRKSRPALYKRPKANGVKRSTVHTI 62
DB 219 SPQRLSTC---QRANSGLDLYAQKNSGVKANG---PRNRRSSYCRMVLSIDIPDLAKHKM 272

QY 63 -----STPOAS--KAISCKGHSISYTLSRNQTV---VEYTHDKDTMFQVGRSTRE-- 109

DB 273 PPAASRTPPYSTNTYTLPCRKSHLSEGATNPDISKMSQGRRAKTTQDVNTGESSEFA 332

QY 110 -SPIDFVVTDTISGSQNTDEAOITOSTISRFAICVCDRNEPTARIFPAQFDSSKNITFL 168

DB 333 DSGIEGATDTDLISRSN--ATNSYS-----PPTGRAF-VGSDSGSS-57 375

QY 169 GEKAAK--WKNPFGHMDGLTTNGVLVNHPRGGFT-----SQGVWRREISVCGDYVTL 220

DB 376 GDRARQGVENFRRELEMSSTNSESLEEGASASDSSGTLSSFG-QSDILLTLTAQGV 434

QY 221 RETRS-----AQQRGLVSESTNVLDGSLDLCGATLLMRTAD---GLFHTPTQKH 270

DB 435 RKGALAAVKNFLVHKQKVKVESATRRKMKNVWSLGGTLFLFETGRSGIDNSPKKA 494

QY 271 EALROEINAAAPQCPVG-----LNTLAFPSINRKEVEBKOPW--AYLSCGHVGYH 321

DB 495 VVWENSIVQAVPEHPKDDFVFCLSNLSGDAFLFQTTSGTELEWMTAHSACAAVARRH 554

QY 322 WGRSDT 328

DB 555 --HKEDT 559

RESULT 4

T03447
dihydrokaempferol 4-reductase (EC 1.1.1.219) A - sorghum

N/Alternate names: dihydroflavonol 4-reductase; NADPH-dependent reductase A1-a

C/Species: Sorghum bicolor (Sorghum)

C/Date: 12-Nov-1999 #sequence_revision 12-Nov-1999 #text_change 09-Jul-2004

A/Accession: T03447

R/Chen, M.; Samiguel, P.; Benmetzen, J.L.

Genetics 148, 435-443, 1998

A/Title: Sequence organization and conservation of Sh2/A1-homologous regions of sorghum

A/Reference number: Z14952; MUID:98133900; PMID:9475753

A/Accession: T03447

A/Status: preliminary; translated from GB/EMBL/DBJ

A/Molecule type: DNA

A/Residues: 1-350 <CHE>

A/Cross-references: UNIPROT:P93776; UNIPARC:UPI00000A710A; EMBL:AF010283; NID:g2735839; I

C/Genetic1:
A/Intons: 52/1; 108/3; 173/3
C/Superfamily: dihydrokaempferol 4-reductase
C/Keywords: oxidoreductase

Query Match 4.4%; Score 100; DB 1; Length 350;

Best Local Similarity 24.1%; Pred. No. 2.1;

Matches 47; Conservative 21; Mismatches 71; Indels 56; Gaps 9;

QY 168 LGKAAKAKNPDGMGLTTNGVLVNHPRGGFTRESQPGVWRREISVCGDYVTLRETRSAQ 227

DB 1 MGEVVAITWATEBG---GAGVKGPPVVTGASGFL---GSLWMLKLLQAGYTVRAF----- 48

QY 228 QRKGLVSESTNVLDGSLDLCGAT---LLW-----RTADGLFHTPTQKH 270

DB 49 -----VRDPAVNVKTYPLDLPLGATERLSLWKADLADBSFPDDATGCTGVFVAATPMD 103

QY 271 EALROEINAAAP-----QCPVG--LNTLAFPSINRKEVEBKOPWAYLSCGHVGYH 318

DB 104 ESKDPENEVKIPVEGEMGMSIMRACKKAGYVRRIVFTSSAGTVNIERRQ-----RPVYD 156

QY 319 YNHWGRSDTEANER 333

DB 157 QDNW---SDVDFCOR 168

RESULT 5

AI0072
probable autotransporter protein YP00587 [imported] - Yersinia pestis (strain CO92)

C/Species: Yersinia pestis
C/Date: 02-Nov-2001 #sequence_revision 02-Nov-2001 #text_change 09-Jul-2004

C/Accession: A10072
 R:Parkhill, J.; Wren, B.W.; Thomson, N.R.; Titball, R.W.; Holden, M.T.G.; Prentice, M.B.;
 demo-Terraga, A.M.; Chillingworth, T.; Cronin, A.; Davies, R.M.; Dave, P.; Dougan, G.;
 11, M.; Rutherford, K.; Simmonds, M.; Skelton, J.; Stevens, K.; Whitehead, S.; Barrett,
 Nature 411, 523-527, 2001
 A/Title: Genome sequence of *Yersinia pestis*, the causative agent of plague.
 A/Reference number: AB0001; MUID:21470413; PMID:11586360
 A/Accession: A10072
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-994 <R>
 A/Cross-references: UNIPROT:Q9P286; UNIPARC:UPI00000DCE70; GB:AL590842; PIDD:CA089444.1;
 C/Genetics:
 A/Genes: YPO0587

Query Match 4.3%; Score 98; DB 2; Length 994;
 Best Local Similarity 22.1%; Pred. No. 13;
 Matches 70; Conservative 53; Mismatches 139; Indels 54; Gaps 17;

QY 13 NKEPVK---YGLVVGYNALPNQD-RGRKSRPLKPKA-NGVKSTVAVISPPA 67
 Db 27 NKEPIATCAAAAFIFSSSVIANPDHGIIVGKSLINKQSAVNALINEGSLVLTDSA 86
 QY 68 SKAISCQGHISIVTISRNTVVEYTHDQDMFQVGRSTESPIDPVVDTISGS--QN 125
 Db 87 SAHTVAVNGSI-FTLKEDSTADIT---SYTGGFFSLSSGSKANINIV---LSGCMLEV 138
 QY 126 TDEAQITOSTISRPACRIVCDRNEPYTARIFAAGFDSKNIPLGEKAAKKNPDGND-- 183
 Db 139 NDDASTETITIS-----SDIEKSTVRLYQDG--SATVTYGDGNDILVSGDSRAEET 189
 QY 184 GLTTNGVLVNHPGPGFTEESQPGVWEISVCGDYVTLRETRSAOQKLYSESTNVLDG 243
 Db 190 HVTKGKGLIYV-----SESQGPTLKNTQTLAGTL-TLKSQVTLGKTEFVSAT-IKTGG 241
 QY 244 SLIDLCATLMTADGLPFTPTQKHEALRQBINARPCQPVLTAPLAPINRKEVVE 303
 Db 242 HLIDNQCQL-----IFNSDKDIYEMKIDGQSLTKENP-LTTLTLSSAG----- 285
 QY 304 EKQPM--AVLSCGHV 317
 Db 286 --DAWVASVYSGETH 299

RESULT 6

C/Accession: C86547
 polymorphic outer membrane protein H family [imported] - Chlamydia pneumoniae (strain
 C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C/Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 09-Jul-2004
 C/Accession: C86547
 R/Shirai, M.; Hirakawa, H.; Kimoto, M.; Tabuchi, M.; Kishi, F.; Ouchi, K.; Shiba, T.; Iw
 Nucleic Acids Res. 28, 2311-2314, 2000
 A/Title: Comparison of whole genome sequences of chlamydia pneumoniae J138.
 A/Reference number: A86491; MUID:20330349; PMID:10871362
 A/Accession: C86547
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-978 <STO>
 A/Cross-references: UNIPROT:Q9Z895; UNIPARC:UPI0000131C83; GB:BA000008; NID:g8978825; PI
 A/Experimental source: strain J138
 C/Genetics:
 A/Genes: pmp_14
 C/Superfamily: Polymorphic membrane protein H family

Query Match 4.2%; Score 95.5; DB 2; Length 978;
 Best Local Similarity 22.1%; Pred. No. 20;
 Matches 47; Conservative 25; Mismatches 74; Indels 67; Gaps 9;

QY 102 FQVGRSTESPIDVVTDTISGSONTDEAQITOSTISRPACRIVCDNE----- 149
 Db 523 FTIGKLAFPDFSLKRPVSAVAGTKNVT-----LTGALVLDHEDVTDLYDMVSLQS 576
 QY 150 PYTARIFAAGFDSKNIPLGEKAAKKNPDGNDGLTTNG-----VLVNHPRG 197

Db 577 PVAIP-----AVFGATVTKGFPDGEIATRSHTGYQKMSYTSRPLIPAPDG 627
 QY 198 GFTEESQP-----GWREISVCGDYVTLRETRSAOQKLYSE-----TNVL 240
 Db 628 GFPGGSPSANTLYAAMNSDGLVRSYTL-----DPERYGEIYVNSLWISFLGNQAFSDIL 683
 QY 241 QDGSLLIDLCATLMTADGL-----FHTPTQKH 269
 Db 684 QDVLIDHPELSI---TAKALGAYVHTPRQGH 713

RESULT 7

C/Accession: B81593
 polymorphic membrane protein H family CP0298 [imported] - Chlamydia pneumoniae (stra
 C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C/Date: 31-Mar-2000 #sequence_revision 31-Mar-2000 #text_change 09-Jul-2004
 C/Accession: B81593
 R/Read, T.D.; Brunham, R.C.; Shen, C.; Gill, S.R.; Heidelberg, J.F.; White, O.; Hickey, H
 C.; Dodson, R.; Gilm, M.; Nelson, W.; DeBoy, R.; Kolonay, J.; McClarty, G.; Salzberg,
 Nucleic Acids Res. 28, 1397-1406, 2000
 A/Title: Genome sequences of Chlamydia trachomatis McpN and Chlamydia pneumoniae AR39.
 A/Reference number: A81500; MUID:20150255; PMID:10684935
 A/Accession: B81593
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-978 <REA>
 A/Cross-references: UNIPROT:Q9Z895; UNIPARC:UPI0000131C83; GB:AE002191; GB:AE002161; NID
 A/Experimental source: strain AR39, HL cells
 C/Genetics:
 A/Genes: CP0298
 C/Superfamily: Polymorphic membrane protein H family

Query Match 4.2%; Score 95.5; DB 2; Length 978;
 Best Local Similarity 22.1%; Pred. No. 20;
 Matches 47; Conservative 25; Mismatches 74; Indels 67; Gaps 9;

QY 102 FQVGRSTESPIDVVTDTISGSONTDEAQITOSTISRPACRIVCDNE----- 149
 Db 523 FTIGKLAFPDFSLKRPVSAVAGTKNVT-----LTGALVLDHEDVTDLYDMVSLQS 576
 QY 150 PYTARIFAAGFDSKNIPLGEKAAKKNPDGNDGLTTNG-----VLVNHPRG 197
 Db 577 PVAIP-----AVFGATVTKGFPDGEIATRSHTGYQKMSYTSRPLIPAPDG 627
 QY 198 GFTEESQP-----GWREISVCGDYVTLRETRSAOQKLYSE-----TNVL 240
 Db 628 GFPGGSPSANTLYAAMNSDGLVRSYTL-----DPERYGEIYVNSLWISFLGNQAFSDIL 683
 QY 241 QDGSLLIDLCATLMTADGL-----FHTPTQKH 269
 Db 684 QDVLIDHPELSI---TAKALGAYVHTPRQGH 713

RESULT 8

G72076
 polymorphic outer membrane protein h family - Chlamydia pneumoniae (strain CML029)
 C/Species: Chlamydia pneumoniae, Chlamydia pneumoniae
 C/Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 09-Jul-2004
 C/Accession: G72076
 R/Kalman, S.; Mitchell, W.; Marathe, R.; Lammel, C.; Fan, J.; Olinger, L.; Grimwood, J.;
 Nature Genet. 21, 385-389, 1999
 A/Title: Comparative genomes of Chlamydia pneumoniae and C. trachomatis.
 A/Reference number: A72000; MUID:99206606; PMID:10192388
 A/Accession: G72076
 A/Status: preliminary
 A/Molecule type: DNA
 A/Residues: 1-978 <ARN>
 A/Cross-references: UNIPROT:Q9Z895; UNIPARC:UPI000046AB; GB:AE001629; GB:AE001363; NID:
 A/Experimental source: strain CML029
 C/Genetics:
 A/Genes: pmp_14
 C/Superfamily: Polymorphic membrane protein H family

Query Match 4.2%; Score 95.5; DB 2; Length 978;
 Best Local Similarity 22.1%; Pred. No. 20;
 Matches 47; Conservative 25; Mismatches 74; Indels 67; Gaps 9;

Query Match 4.2%; Score 95.5; DB 2; Length 978;
Best Local Similarity 22.1%; Pred. No. 20;
Matches 47; Conservative 25; Mismatches 74; Indels 67; Gaps 9;

QY 102 FQVGRSTESPIDVVTDTISGSONTDEAQTOSTISRACRYCDRNE----- 149
DB 523 FTIGKIAFDPSFLKEDFVSASVNACTKNT-----LTGALVLDHVDVLDYDWSLQS 576
QY 150 PYTARIFAAGFDSKKNIFLGEKAAKMKNDGMDGLTNG-----VLVNHPRG 197
DB 577 PVAIPF-----AVFKATATVTKGPPDGEIATPSHHYQCKMSYMSRPLIPADG 627
QY 198 GFTEESOP-----GVWRISVCGDVYTLRETSAQORGLVSEB-----TNVL 240
DB 628 GFEGGSPSANTLYXAVNNSSTLVRSYTL-----DPERYGVISNSLMISFLGNQARSDDL 683
QY 241 QDGSLLIDLCATLLMRTADL-----FHTPTQKH 269
DB 684 QDVLLIDHPLSLI---TAKALGAVERTPRQGH 713

RESULT 9

C81804
alanine-tRNA ligase (EC 6.1.1.7) NMA1788 [Imported] - Neisseria meningitidis (strain Z24
C/Species: Neisseria meningitidis
C/Date: 05-May-2000 #sequence_revision 05-May-2000 #text_change 09-Jul-2004
C/Accession: C81804
R/Parikh, J.; Achman, M.; James, K.D.; Bentley, S.D.; Churcher, C.; Klee, S.R.; Morel
R.; Holtz, S.; Uggelø, K.; Leather, S.; Moutle, S.; Mungall, K.; Quail, M.A.; Rajandream,
Nature 404, 502-506, 2000
A/Title: Complete DNA sequence of a serogroup A strain of Neisseria meningitidis Z2491.
A/Reference number: A81775; MUID:20222556; PMID:10761919
A/Accession: C81804
A/Status: preliminary
A/Molecule type: DNA
A/Residues: 1-874 <PAR>
A/Cross-references: UNIPROT:Q9JTG4; UNIPARC:UPI0000136315; GB:AL162757; GB:AL179599; NID
A/Experimental source: serogroup A, strain Z2491
C/Genetics:
A/Gene: alag; NMA1788
C/Superfamily: alanyl-tRNA ligase
C/Keywords: ligase

Query Match 4.0%; Score 92.5; DB 2; Length 874;

Best Local Similarity 23.0%; Pred. No. 30; Mismatches 127; Indels 77; Gaps 15;

QY 66 QASKAISCKQHSISYTLNRQT---VVVEYTHDKOTDMFQVGRSTESPIDFVVTDTISG 122
DB 441 KANAQLEPYDQDTEFKYSERQTESKYLALYKQSEVNEINEDSGAVVIDFTPTVAESG 500
QY 123 SQNTDEAQTOSTISRACRYCDRNEPYTARIFA--GPDSSKNIFGE-----K 171
DB 501 GQVGDVGYIRFSS--NRFEVR---DTQKIRAAVFGQFQVOTSGRLKGDVTAQVDEIR 555
QY 172 AAKMKND-----GHMDGLTNGVYV-----MHRGFTSE----- 202
DB 556 NANNRNNSATLHMKALRDVGGHVE---QKSLVTAESTRPDISHOAVTAEBIAEVER 612
QY 203 -SOPGVREISVCGDVYTLRETSAQORGLV-----ESETVVLQDGSV-IDLCATLL 254
DB 613 RVNEAVLVANVAVNAATISMED---AQKTGAMMLEGKYGGEVNLQWCGFSTELCGSTHV 669
QY 255 WRPAD-GLFHTPTQKHIEALRQETINARPOCPVGLNTLAPPSINRKEVE-----E 304
DB 670 SRFGIDGLFRTISEGCIAGVRIRIET---TGLNALKWAQEQERLVKDIIAETKAQTE 724
QY 305 KQPMVYLSGHHV 317
DB 725 KDVLAKEIQAGAAH 737

RESULT 10

S42373
hypothetical protein T20G5.3 - Caenorhabditis elegans
C/Species: Caenorhabditis elegans
C/Date: 07-Oct-1994 #sequence_revision 10-Nov-1995 #text_change 09-May-2004
C/Accession: S42373
R/Smith, A.
submitted to the EMBL Data Library, March 1994
A/Reference number: S42368
A/Accession: S42373
A/Molecule type: DNA
A/Residues: 1-3051 <SMI>
A/Cross-references: UNIPARC:UPI000017B8C8; EMBL:Z30423; NID:G458479; PID:G458485
C/Genetics:
A/Intons: 96/1; 166/3; 201/2; 355/1; 404/1; 427/2; 794/1; 892/1; 952/2; 1041/1; 1146/1;
F/512-679/Domain: von Willebrand factor type A repeat homology <WAL>
F/754-793/Domain: fibronectin type II repeat homology <ZFI>
F/1201-1244/Domain: EGF homology <EGF>

Query Match 4.0%; Score 92; DB 2; Length 3051;

Best Local Similarity 21.6%; Pred. No. 1,8e+02;
Matches 94; Conservative 44; Mismatches 156; Indels 142; Gaps 24;

QY 4 PQG---EEHCAPNK-----EP---VKYGLVVLGYNGALPNGDRGRKRSREPLVYR 48
DB 1687 PGRVCSAPRCVCPNHDSSAAVCEPLGGMKYCCVCIOGYVDQSPGSKGRVCVRNACHD 1746
QY 49 PRKNGVYKPTSTVAVISTPQASKAISCKQHSISYTLNRQTVVEYTH-DKQTDMPQVGRS 107
DB 1747 PRLN-----TC-SRNAICTDEPRGYRCECKGFMDRSPSSQGRV 1786
QY 108 TESP-----IDFVVTDTISGSONTDEAQTOSTISRACRYCDRNE 149
DB 1787 CEPPPPSPRRPHPCDPERNDCHPACTGATAGS-----YTCBGLSGVADRSPPRRK 1841
QY 150 PYTARIFA--AGFDSKKNIFLGEKAAKMKNDGMDGLT--NG--VLVNHPRGFTES 203
DB 1842 PGRLCVLTSPVCLDPRQN-----DCHAAALICEVNGPEKTYCKCRDGYTDES 1888
QY 204 QPQVMEISVCGDVYTLRETSAQORGLVSESTNVLQDGSLLDC-----CATLLMRA 258
DB 1889 PDLPRRPRGIC-----KGLNECLDRSLNCHSLAVCKDLPNGT----- 1928
QY 259 DGLFHTPTQKHIEALRQETINARPO--CPVGLNTLAFPSINR---KEVVEKQPMAYLS 312
DB 1929 -----CCCPINAKQSPDRPKRGKICSLSVNECANPSLNCSSAFADPDENGY---R 1978
QY 313 CGHNGYH-----NMGRSDTEANR-----RECPMCRVGPYVPLMIGCEAGFYV 357
DB 1979 CRCRNGYHDDDPRAHPGRCSFMINECDSSNLNDCBNANCIDTAGGYD---CAKKA-PYR 2034
QY 358 DAGPPTHAFTPCGHVC 373
DB 2035 DEGPPO---SP-GRIC 2046

RESULT 11

T40476
hypothetical protein SPBC4B4.04 - fission yeast (Schizosaccharomyces pombe)
C/Species: Schizosaccharomyces pombe
C/Date: 03-Dec-1999 #sequence_revision 03-Dec-1999 #text_change 09-Jul-2004
C/Accession: T40476
R/Beck, A.; Reinhardt, R.; Lyne, M.; Wood, V.; Rajandream, M.A.; Barrell, B.G.
submitted to the EMBL Data Library, May 1997
A/Reference number: Z21932
A/Accession: T40476
A/Status: preliminary; translated from GB/EMBL/DBJ
A/Molecule type: DNA
A/Residues: 1-576 <BEC>
A/Cross-references: UNIPROT:O74965; UNIPARC:UPI000006A650; EMBL:AL023706; PIDN:CAA19284.1
A/Experimental source: strain 972h-; cosmid c4B4
C/Genetics:
A/Gene: SPBC4B4.04

A:Map position: 2
A:Insertions: 9/1
C:Superfamily: translation initiation factor eIF-2A

Query Match 4.0%; Score 91.5; DB 2; Length 576;
Best Local Similarity 20.6%; Pred. No. 20;
Matches 56; Conservative 35; Mismatches 98; Indels 83; Gaps 12;

39 RKSRLALYKRPKNGV-----KPTVAVISPPQASKAISCKGCHISITSLRNQTV 90
3 OKQFA-YRSSKSLIGLVNASENATSPKFAISSPANNACYSPPNGK----- 47
91 VERTHDKDTMFQGRSTESPIDFVVTDTISGQNTDEAQTOSTISRPACRIYCDRNEP 150
48 -----LFAVATITQ-----VINDTSSG-----ALITDLP----- 72
151 YTAIRPAGFDSCKNIFLGEKAAKKNPDGHMDGLTTNGVLVMPHPRG-----FTESQP 205
73 -AANTYELGFSP-----LGKYLSTWERPGKADGTQPKNNKVMNTERGQVFSFVQRNQT 126
206 GWRREISVCGDVTTLARETSQAQGRKLVSEETNVLQDGLDLCATLTLMTADGLFHTP 265
127 G-WNL-----QYTCESLAA-----RLVTNEVHFYETGNNSKGIATLREGISDFALSP 175
266 TOKHIEALROEINAAAPQCPVGLNTLAFPSIN 297
176 GONHAAV-----FTEPKGAPASVATTSIENFN 204

RESULT 12
A:2017
hypothetical protein all1695 [imported] - Nostoc sp. (strain PCC 7120)

C:Species: Nostoc sp. PCC 7120
A:Note: Nostoc sp. strain PCC 7120 is a synonym of Anabaena sp. strain PCC 7120
C:Date: 14-Dec-2001 #sequence_revision 14-Dec-2001 #text_change 09-Jul-2004
C:Accession: A12017
R:Kaneho, T.; Nakamura, Y.; Wolk, C.P.; Kuritz, T.; Sasamoto, S.; Watanabe, A.; Iriuchihara, N.; Shimpo, S.; Sugimoto, M.; Takazawa, M.; Yamada, M.; Yasuda, M.; Tabata, S.
DNA Ref. 8, 205-213, 2001
A:Title: Complete Genomic Sequence of the Filamentous Nitrogen-fixing Cyanobacterium Anabaena sp. PCC 7120
A:Reference number: AB1807; MUID:21595285; PMID:11759840
A:Accession: A12017
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-1449 <RUR>
A:Cross-references: UNIPROT:Q8YWC0; UNIPARC:UP100000CE194; GB:BA000019; PIDN:BA078061.1;
A:Experimental source: strain PCC 7120
C:Genetics:
A:Gene: all1695

Query Match 4.0%; Score 91.5; DB 2; Length 1449;
Best Local Similarity 22.5%; Pred. No. 72;
Matches 74; Conservative 56; Mismatches 126; Indels 73; Gaps 19;

12 PNEKPVKYGELVVLGNGALPNDGRGRKRPALYKPKANGVCPST---VHVISTQAS 68
44 PNVAVAVY-----KNQKITVOLEKNSQNFAYTLA-HIGVKKETLVALLYERSPDVI 94
69 KAI-----SKGQH-----SISYTLISNQTWVEYTHDKDTMDFQVGRSTESPIDFV---VTDI 119
95 IAILAIKAGAYPLDISAPLELEHTI-----KNAEAF-IITTESQNTYIEKTTDE 147
120 ISSQNTDEAQTOSTISRPACRIYCDRNEPYTAIRPAGFDSCKNIFLGEKAAKKNPD 179
148 IQTICIDNQNLTDFSDISLCEVTANLAVV---MTSG-----ST 187
180 GHMDGLTT---NGVLVMPHPRGFTESQPGWREI-SVCGDVTTLARETSQAQGRKLVSE 236
188 GKPKGVCEHGRVRLVKNNTYANFSDEVTILQASIAFPAATYEIWAALLNGKLVMP 247
237 TTV--LQD-GSLDLCATLTLMTADGLFHTPTQ---KHIEALROEINAAAPQCPVGLNT 290
248 INISLQEIEMALQYHTVITLMTA-GLENLMEVEQIEHLKSLQQL-----AGGDV 298

291 LAFPSINRKEVEKOPMAYLSCGHVGY 319
299 LSVYHVS-K-VIEELP-----NCQLINGX 320

RESULT 13
B33830
cation efflux system membrane protein czcb - Alcaligenes eutrophus

C:Species: Alcaligenes eutrophus
C:Date: 23-Mar-1990 #sequence_revision 18-Sep-1992 #text_change 09-Jul-2004
C:Accession: B33830
R:Nies, D.H.; Nies, A.; Chu, L.; Silver, S.
Proc. Natl. Acad. Sci. U.S.A. 86, 7351-7355, 1989
A:Title: Expression and nucleotide sequence of a plasmid-determined divalent cation efflux system membrane protein; lipoyl/biotin-binding homology #status atypical <LPB>
A:Reference number: A33830; MUID:90017477; PMID:2678100
A:Accession: B33830
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-520 <NIB>
A:Cross-references: UNIPROT:P13510; UNIPARC:UP10000128D83
C:Superfamily: cation efflux system membrane protein; lipoyl/biotin-binding homology
F;217-260,333-362/Domain: lipoyl/biotin-binding homology #status atypical <LPB>

Query Match 4.0%; Score 90.5; DB 2; Length 520;
Best Local Similarity 23.3%; Pred. No. 21;
Matches 53; Conservative 23; Mismatches 94; Indels 57; Gaps 8;

94 THDKDTMFQVGRSTESPIDFVVTDTISGQNTDEAQTOSTISRPACRIYCDRNEP--- 150
277 TYREKKLWEGKISAQ-----DYSARNALQEAQISQVNAQQLTAIGANSSTALN 329
151 -YTAIRPAGFDSCKNIFLGEKAAKKNPDGHMDGLTTNGVLVMPHPRGFTESQPGWVR 209
330 RYELRAPFAGMIVEKHSIGEAVALDANAV-----FTLSDLSSVMA 369
210 EISVCG-DVYTLA--ETRSQAQGRKLVSEETNVLQDGL-----IDLQATLW 255
370 EFVSAKDVYERVIGKASINSNSDVKADGVSYVGSLLGEQRTAKARVTLTNPMW 429
256 RTADGLFHTPTQKHIEALROEINAAAPQCPVGLNTLAFPSINKEV 302
430 R--PGLFVT-----VDVFGADVEVPVAVKTAVDVNGESV 464

RESULT 14
P90914
hypothetical protein Ec82286 [imported] - Escherichia coli (strain O157:H7, substrain RIN)

C:Species: Escherichia coli
C:Date: 18-Jul-2001 #sequence_revision 18-Jul-2001 #text_change 09-Jul-2004
C:Accession: P90914
R:Hayashi, T.; Makino, K.; Ohnishi, M.; Kurokawa, K.; Ishii, K.; Yokoyama, K.; Han, C.G.; Gasawara, N.; Yasunaga, T.; Kuhara, S.; Shiba, T.; Hatoto, M.; Shinagawa, H.
DNA Ref. 8, 11-22, 2001
A:Title: Complete genome sequence of enterohemorrhagic Escherichia coli O157:H7 and genome
A:Reference number: A9629; MUID:21156231; PMID:11258796
A:Accession: P90914
A:Status: preliminary
A:Molecule type: DNA
A:Residues: 1-823 <HAY>
A:Cross-references: UNIPROT:Q8X280; UNIPARC:UP100000D0F1F; GB:BA000007; PIDN:BA035709.1;
A:Experimental source: strain O157:H7, substrain RIMD 0509952
C:Genetics:
A:Gene: Ec82286

Query Match 4.0%; Score 90.5; DB 2; Length 823;
Best Local Similarity 21.2%; Pred. No. 40;
Matches 77; Conservative 39; Mismatches 126; Indels 121; Gaps 20;

10 CA--PNEKPVKYGELVVLGNGALPNDGRGRGRF---ALYKRPANGVCPSTHVIST 64
7 CAALPDEQALKEGAVAVA--TIAEGDERRAAKHMQELFHTYPAADQ----- 54

QY 65 PQASKAISCCKGSHISYTLSTSRNOTVVEYTHDKDTDFQVGRSTESPIDVVTDTISGQ 124
 Db 55 --AYKFLVCDKXGIRPRA-----LDSDAEVWQGNRNDDEASASVVPVETESDPM 102
 QY 125 NTDEAOITOSTISRFACRIYCDRNEPYTARIFAAGFDSKNI-----FLGKQAK 174
 Db 103 N-----VTFDMILAPFVQNAVWVKPFCPTCENITVDWVISAQELQDEMMAT 145
 QY 175 WKNPDCG-----MDGLTTNGV---LVWHPRGFTBESQPGV-NREISVCDVYTLARTRS 225
 Db 146 F---DGIYVALKMPEVNMVPELKIHAIGWVHKKCPAKMPEIQAEKRIWKKREGE 202
 QY 226 AOCGRKLVESETVWLODGLIDCGATLWRTADGLFHTPTOKHIEALROEINAAAPQC 285
 Db 203 RKAGKKT-----SVVDLARA-----RANQCHTSTGKIN-----P 234
 QY 286 VGLNTLAPSPINRKEVEEKQPM-----AY-----LSCGHVHG-YHNWGHRSDEAN 331
 Db 235 V-----IAAIHR-----EYKQWTWKTLDDELALVAMPQDVADAGNIDGSIHRMA-KNEVYDN 283
 QY 332 ERE 334
 Db 284 GRE 286

RESULT 15
 NBHDS
 Complement factor H precursor, short splice form [validated] - human
 N/Alternate names: complement factor H-related protein; complement protein H
 C/Species: Homo sapiens (hmn)
 C/Date: 31-Dec-1993 #sequence revision 23-Feb-1996 #text change 09-Jul-2004
 C/Accession: S03013; B60238; A27877; A61103; A26505; S10479
 R/Rtrophe, J.; Day, A.J.; Harris, T.O.R.; Sim, R.B.
 Biochem. J. 249, 593-602, 1988
 A/Title: The complete amino acid sequence of human complement factor H.
 A/Reference number: S00254; MUID:88134059; PMID:2963625
 A/Accession: S03013
 A/Molecule type: mRNA
 A/Residues: 1-449 <RP>
 A/Cross-references: UNIPROT:P08603; UNIPARC:UPI000002635; EMBL:X07523; EMBL:Y00716; NID:402-Tyr was also found
 A/Note: part of this sequence, including the amino end of the mature protein was conflict
 R/Estaller, C.; Schwaeble, W.; Dietrich, M.; Weiss, E.H.
 Eur. J. Immunol. 21, 799-802, 1991
 A/Title: Human complement factor H: two factor H proteins are derived from alternatively
 A/Reference number: A60238; MUID:91184292; PMID:1826264
 A/Accession: B60238
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 1-33; 434-449 <EST>
 A/Cross-references: UNIPARC:UPI000017428; UNIPARC:UPI000017429
 A/Note: only portions of this 1.8 kilobase mRNA were sequenced
 R/Schulz, T.F.; Schwaeble, W.; Stanley, K.K.; Weiss, E.; Dietrich, M.P.
 Eur. J. Immunol. 16, 1351-1355, 1986
 A/Title: Human complement factor H: isolation of cDNA clones and partial cDNA sequence c
 A/Reference number: A27877; MUID:87054207; PMID:2946589
 A/Accession: A27877
 A/Molecule type: mRNA
 A/Residues: 1-55; 401, 'Y', 403-449 <SCH>
 A/Cross-references: UNIPARC:UPI000016A08; GB:X04697; NID:931991; PIDN:CAB41739.1; PID:8
 A/Note: an additional nucleotide present within the codon for Glu-310 was thought to be
 R/Schaeble, W.; Zwirner, J.; Schulz, T.F.; Linke, R.P.; Dietrich, M.P.; Weiss, E.H.
 Eur. J. Immunol. 17, 1485-1489, 1987
 A/Title: Human complement factor H: expression of a second truncated gene product c
 A/Reference number: A61103; MUID:88055295; PMID:2445583
 A/Accession: A61103
 A/Status: not compared with conceptual translation
 A/Molecule type: mRNA
 A/Residues: 27-76 <SC2>
 A/Cross-references: UNIPARC:UPI00001742EA
 A/Note: this is the partial sequence of an alternatively spliced 1.8 kilobase mRNA that is
 R/Sim, R.B.; Discipio, R.G.

Biochem. J. 205, 285-293, 1982
 A/Title: Purification and structural studies on the complement-system control protein bet
 A/Reference number: A26505; MUID:83048213; PMID:6215918
 A/Accession: A26505
 A/Molecule type: protein
 A/Residues: 19-20, 'O', 22-29, 'V', 31-33, 'Q', 35 <SIM>
 A/Cross-references: UNIPARC:UPI00001742B5
 R/Bartlow, P.N.; Norman, D.G.; Steinkasserer, A.; Horne, T.J.; Pearce, J.; Driscoll, P.C.;
 Biochemistry 31, 3626-3634, 1992
 A/Title: Solution structure of the fifth repeat of factor H: A second example of the com
 A/Reference number: A44551; MUID:92232649; PMID:1533152
 A/Contents: annotation; NMR structure determination, residues 264-292
 R/Kristensen, T.; Wetzel, R.A.; Tack, B.F.
 J. Immunol. 136, 3407-3411, 1986
 A/Title: Structural analysis of human complement protein H: homology with C4b binding pr
 A/Reference number: S10479; MUID:86169701; PMID:2537845
 A/Accession: S10479
 A/Molecule type: mRNA
 A/Residues: 226-401, 'Y', 403-449 <KRI>
 A/Cross-references: UNIPARC:UPI000016A6C8; GB:M12383; NID:9180472; PIDN:AAA52013.1; PID:5
 C/Comment: Factor H has also been found bound to cell membranes in an unknown manner. How
 C/Comment: Alternative transcripts of 4.3, 1.8, and 1.4 kilobases are expressed in liver.
 C/Genetics: <HF1>
 A/Gene: GDB:HF1; HF
 A/Cross-references: GDB:120041; OMIM:134370
 A/Map position: 1q32-1q32
 C/Genetics: <HF2>
 A/Gene: GDB:HF2; HF
 A/Cross-references: GDB:129095
 A/Map position: 1q32-1q32
 A/Note: the correspondence between the two loci and the sequences indicated is unclear; f
 C/Function:
 A/Description: a cofactor in the inactivation of C3b by serine proteinase I; also increas
 he alternative complement pathway
 A/Pathway: complement alternate pathway
 C/Keyword: complement factor H; complement factor H repeat homology
 C/Keywords: alternative splicing; complement alternate pathway; glycoprotein; plasma
 F:1-18/Domain: signal sequence #status predicted <Sig>
 F:19-449/Product: complement factor H, short splice form #status experimental <MAT>
 F:21-80/Domain: complement factor H repeat homology <FH01>
 F:85-141/Domain: complement factor H repeat homology <FH02>
 F:146-205/Domain: complement factor H repeat homology <FH03>
 F:210-262/Domain: complement factor H repeat homology <FH04>
 F:246-248/Region: cell attachment (R-G-D) motif
 F:267-320/Domain: complement factor H repeat homology <FH05>
 F:325-385/Domain: complement factor H repeat homology <FH06>
 F:389-442/Domain: complement factor H repeat homology <FH07>
 F:21-66, 52-80, 85-129, 114-141, 146-192, 178-205, 210-251, 237-262, 267-309, 294-320, 325-374, 357-
 F:217/Binding site: carbohydrate (Asn) (covalent) #status absent

Query Match 3.9%; Score 90; DB 1; Length 449;
 Best Local Similarity 20.1%; Pred. No. 19;
 Matches 67; Conservative 37; Mismatches 124; Indels 106; Gaps 15;

QY 117 TDTISGQNTDEAOITOSTISRFACRIYCDRNEPYTARIFAAGFDSKNIPLGEKAKY- 175
 Db 30 TELTGSWS--DPTYREGQALYKCR-----EYRLSGNVIMYCRGENV 72
 QY 176 -----KNPDGNDGLTTNGVLVWHPRG-----FTBE 202
 Db 73 ALNPLRKCQGRPCGH--PGDTPFGTFTL--TGVNFEYGVAVYTCNEGYYQLLGEINVRG 129
 QY 203 SQGVKREISVCGDYVTLAETRSQOGRGLVSEETN-----VLQDGLSIDLC 249
 Db 130 DTDGWTNDIPICGVVCLPVT--APENGKIVSSAMEPDBEYHGAQVRFVCSNGYKIE-- 185
 QY 250 GATILWRTADGLFHTPTOKHIEALROEINAAAPQCQVGLNTLAPSPINRKEVEEKQPM 309
 Db 186 GDEKMGCSDDGFGFSKKPKCV-----EICKSPDVING-----SFSQKIYKEMERFQ 234
 QY 310 YLSCGHVHGYNWGHRSDEANERECPMKRTV--GPYP-----LMTG 350
 Db 235 Y-KCMNGEYSEKGDVAVCTESGWRPLPSCBEKSCDNPYIPNGDYSPRLIKHRTGDEITVQ 293

OY 351 CEAGFYVDAGPPTTHAFTPCGHV----CSEKSAKY 380
Db 294 CRNGFTPATRGNTAKCTSTGMIAPRCTLPCDY 327

Search completed: March 1, 2006, 21:00:48
Job time : 42 secs

This Page Blank (uspto)

this page blank (uspto)

GenCore version 5.1.7
Copyright (c) 1993 - 2006 Bioceleration Ltd.

OM protein - protein search, using sw model

Run on: March 1, 2006, 20:52:46 ; Search time 185 Seconds
(without alignments)
997.509 Million cell updates/sec

Title: US-10-041-030-4

Perfect score: 2290
Sequence: 1 MFSPGGEHCAFNKPKYK.....ATQLVGRNCIKLIFGSPID 420

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 2443163 seqs, 439378781 residues

Total number of hits satisfying chosen parameters: 2443163

Minimum DB seq length: 0
Maximum DB seq length: 200000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database :

A_Geneseq_21:*
1: geneseqp1980s:*
2: geneseqp1980s:*
3: geneseqp2000s:*
4: geneseqp2001s:*
5: geneseqp2002s:*
6: geneseqp2003as:*
7: geneseqp2003bs:*
8: geneseqp2004s:*
9: geneseqp2005s:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

| Result No. | Score | Query Match | Length | ID | Description |
|------------|--------|-------------|--------|----|--------------------|
| 1 | 2290 | 100.0 | 420 | 5 | ABB78077 Amino aci |
| 2 | 2290 | 100.0 | 420 | 5 | AAO17522 Human M33 |
| 3 | 2290 | 100.0 | 420 | 6 | ABR41083 Human MAP |
| 4 | 2290 | 100.0 | 420 | 7 | ADC31180 Human nov |
| 5 | 2283 | 99.7 | 420 | 5 | ABB07922 Human pel |
| 6 | 2283 | 99.7 | 420 | 5 | ADP48671 Human pel |
| 7 | 2175.5 | 95.0 | 419 | 5 | ABB07921 Murine pe |
| 8 | 2175.5 | 95.0 | 419 | 8 | ADP48669 Mouse pel |
| 9 | 2111.5 | 92.2 | 419 | 6 | ABR41082 Mouse MAP |
| 10 | 1917 | 83.7 | 418 | 5 | ABB07920 Human pel |
| 11 | 1917 | 83.7 | 418 | 5 | AAO17518 Human M30 |
| 12 | 1917 | 83.7 | 418 | 6 | ABU00017 Human nov |
| 13 | 1917 | 83.7 | 418 | 8 | ADF83103 Human pel |
| 14 | 1917 | 83.7 | 418 | 8 | ADP44005 Amino aci |
| 15 | 1917 | 83.7 | 418 | 8 | ADP48667 Human pel |
| 16 | 1917 | 83.7 | 418 | 8 | ADP56420 Human PRO |
| 17 | 1917 | 83.7 | 418 | 8 | ADP24929 PRO polyp |
| 18 | 1917 | 83.7 | 418 | 9 | ADY26614 PRO polyp |
| 19 | 1917 | 83.7 | 418 | 9 | ADY17390 PRO polyp |
| 20 | 1916 | 83.7 | 418 | 5 | AAO17515 Rat M30 p |
| 21 | 1913 | 83.5 | 418 | 5 | ABB78078 Amino aci |
| 22 | 1912 | 83.5 | 418 | 5 | ABB07919 Murine pe |
| 23 | 1912 | 83.5 | 418 | 8 | ADP48665 Mouse pel |
| 24 | 1910 | 83.4 | 418 | 5 | AAO17519 Human M30 |

| | | | | | |
|----|--------|------|-----|---|--------------------|
| 25 | 1642 | 71.7 | 445 | 5 | ABB07923 Human pel |
| 26 | 1642 | 71.7 | 445 | 7 | ADD71147 Human int |
| 27 | 1642 | 71.7 | 445 | 8 | ADP48675 Human pel |
| 28 | 1640 | 71.6 | 445 | 5 | AAO17521 Human M31 |
| 29 | 1633 | 71.3 | 445 | 7 | ADC31807 Human nov |
| 30 | 1632 | 71.3 | 445 | 5 | AAO17520 Murine M3 |
| 31 | 1582.5 | 69.1 | 415 | 8 | ADP44009 Amino aci |
| 32 | 1492 | 65.2 | 320 | 5 | AAO17516 Human M30 |
| 33 | 1492 | 65.2 | 320 | 5 | AAO17517 Human M30 |
| 34 | 1482.5 | 64.7 | 406 | 7 | ADP85774 Human pro |
| 35 | 1482 | 64.7 | 320 | 4 | AAO17518 Human pro |
| 36 | 1482 | 64.7 | 320 | 4 | AAO17519 Human pro |
| 37 | 1302 | 56.9 | 424 | 4 | AAO17520 Human pro |
| 38 | 1302 | 56.9 | 424 | 5 | AAO17521 Human pro |
| 39 | 1302 | 56.9 | 424 | 8 | ADP48676 Human pro |
| 40 | 1300 | 56.8 | 320 | 4 | AAO17522 Human pro |
| 41 | 1300 | 56.8 | 320 | 8 | ADL31994 Human pro |
| 42 | 1235 | 53.9 | 455 | 5 | ABB07925 Amino aci |
| 43 | 1235 | 53.9 | 455 | 8 | ADP48677 Human pro |
| 44 | 1102 | 48.1 | 196 | 7 | ADC32882 Human pro |
| 45 | 1077 | 47.0 | 241 | 4 | ABG04494 Novel hum |

ALIGNMENTS

RESULT 1
ID ABB78077 standard; protein; 420 AA.
XX
AC ABB78077;
DT 05-NOV-2002 (first entry)
XX
XX Amino acid sequence of human Pellino 2.
DE
XX Pellino 2; cancer; Pellino 1; cancer treatment; epithelial cancer;
KW gastrointestinal tract cancer.
XX
XX Homo sapiens.
OS
XX
XX WO200259641-A2.
XX
XX 01-PAG-2002.
XX
XX 28-DEC-2001; 2001MO-US051368.
XX
XX 02-JUN-2001; 2001US-0259502P.
XX
XX (TULIA-) TULARIK INC.
XX
XX Powers S, Mu D, Xiang P, Peng Y;
PI
XX WPI: 2002-619185/66.
XX
XX N-PSDB; ABQ78319.
XX
XX Detecting cancer cells in mammalian sample, useful for identifying
PT inhibitors for treating cancer e.g. epithelial cancer, comprises
PT detecting an overexpression of, or increase in copy number of genes
PT encoding, Pellino 1 and Pellino 2.
XX
XX Claim 35; Page 57-61; 69pp; English.
XX
XX The present sequence represents Pellino 2. The specification describes a
XX method for detecting cancer cells in biological sample from a mammal. The
XX method comprises detecting an overexpression of, or increase in copy
XX number of genes encoding, polypeptides Pellino 1 or Pellino 2. The method
XX is useful in detecting cancer or propensity to develop cancer, monitoring
XX the efficacy of cancer treatment, identifying inhibitors of Pellino 1 and
XX 2, inhibiting the expression and/or activity of Pellino 1 and 2 in cancer
XX cells, and treating cancer or inhibiting proliferation of cancer. The
XX cancer can be epithelial cancer, such as lung, colon, ovarian, breast,
XX prostate, kidney, stomach, bladder, or any cancer of the gastrointestinal

CC tract
XX
SQ Sequence 420 AA;

Query Match 100.0%; Score 2290; DB 5; Length 420;
Best Local Similarity 100.0%; Pred. No. 2.8e-226;
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSRQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRPALTYPKRPANGVKPSTVH 60
DB 1 MFSRQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRPALTYPKRPANGVKPSTVH 60
QY 61 VISTPQASKAISCKGHSISYTLISRNQTVVEYTHDKDTMPQVGRSTSPIDFVVTDTI 120
DB 61 VISTPQASKAISCKGHSISYTLISRNQTVVEYTHDKDTMPQVGRSTSPIDFVVTDTI 120
QY 121 SSGQNTDEAQTOSTISRFACRIYCDRNEPYTARIFAAGFDSKNIPLGEKAKMKNPDG 180
DB 121 SSGQNTDEAQTOSTISRFACRIYCDRNEPYTARIFAAGFDSKNIPLGEKAKMKNPDG 180
QY 181 HMDGLTNGVLVMPHPRGFTBESQPGVWREISVCGDYTLRRTSAQORGLVSESTNVL 240
DB 181 HMDGLTNGVLVMPHPRGFTBESQPGVWREISVCGDYTLRRTSAQORGLVSESTNVL 240
QY 241 QDGSLLDLCGATLLMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
DB 241 QDGSLLDLCGATLLMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
QY 301 VVEKQPMAYLSCGHVGHYNNWGHRSOTENARECPMCRITVGPVPLMTGCEAGFYVDAG 360
DB 301 VVEKQPMAYLSCGHVGHYNNWGHRSOTENARECPMCRITVGPVPLMTGCEAGFYVDAG 360
QY 361 PPTHAFPCGHVSEKSAKYWSQIPLPHGTHFAHACPCATQLVGEONCIKLIPOGPID 420
DB 361 PPTHAFPCGHVSEKSAKYWSQIPLPHGTHFAHACPCATQLVGEONCIKLIPOGPID 420

RESULT 2

AA017522
ID AA017522 standard; protein; 420 AA.

XX
AC AA017522;

DT 19-JUL-2002 (first entry)

XX
DE Human M33 protein SEQ ID NO: 16.

XX
KW Neurodegenerative disease; M30; M31; M32; M33; stroke;
KW fragile X syndrome; Huntington's disease; Parkinson's disease;
KW Alzheimer's disease; multiple sclerosis; ovarian cancer;
KW neurodegeneration; immune disorder; autoimmune disease; allergy;
KW infection; leukemia; inflammation; neuroprotective; cerebroprotective;
KW immunosuppressive; cytotoxic; motropic; antiparkinsonian; antiallergic;
KW virucide; antiinflammatory.

XX
OS Homo sapiens.

XX
PN W0200221138-A2.

XX
PD 14-MAR-2002.

XX
PF 07-SEP-2001; 2001MO-EP010366.

XX
PR 07-SEP-2000; 2000US-00657479.

XX
PA (AXAR-) AXARON BIOSCIENCE AG.

XX
PI Schneider A, Hienisch H, Rosner M, Klugmann M, Naim J;

XX
PI Eisenhardt G, Kuner R, Lanahan A, Worley P, Spielvogel D, Scheek S;

XX
DR WPI; 2002-292287/33.

XX
DR N-PSDB; AAL46323.

XX

PT Diagnosis of neurodegenerative disease comprises detecting level of M30-
PT family proteins.

XX
PS Claim 2; Page 116-117; 130pp; German.

XX
CC The present invention relates to a method of diagnosing neurodegenerative
CC diseases, comprising determining the concentration of a protein in a body
CC sample, where the protein may be M30 or a variant thereof, M31, M32 or
CC M33. The method is used to diagnose neurodegenerative diseases,
CC particularly stroke but also e.g. fragile X syndrome, Huntington's,
CC Parkinson's and Alzheimer's diseases, multiple sclerosis etc. Also
CC overexpression of M31 can be used for diagnosis of carcinoma and sarcoma,
CC especially ovarian cancer. The proteins can be used to identify specific
CC ligands, potentially useful for treating neurodegeneration, immune-system
CC disorders (e.g. autoimmune diseases, allergy, viral infection, leukemia,
CC inflammation etc.), carcinoma and sarcoma. Inhibitors of the interaction
CC between the proteins and the protein kinase IRAK-1 can be used to treat
CC neurodegeneration. The present sequence is a protein used in the method
CC of the invention

XX
SQ Sequence 420 AA;

Query Match 100.0%; Score 2290; DB 5; Length 420;
Best Local Similarity 100.0%; Pred. No. 2.8e-226;
Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MFSRQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRPALTYPKRPANGVKPSTVH 60
DB 1 MFSRQGEHCAPNKEPVKYGELVVLGYNGALPNDGRGRKSRPALTYPKRPANGVKPSTVH 60
QY 61 VISTPQASKAISCKGHSISYTLISRNQTVVEYTHDKDTMPQVGRSTSPIDFVVTDTI 120
DB 61 VISTPQASKAISCKGHSISYTLISRNQTVVEYTHDKDTMPQVGRSTSPIDFVVTDTI 120
QY 121 SSGQNTDEAQTOSTISRFACRIYCDRNEPYTARIFAAGFDSKNIPLGEKAKMKNPDG 180
DB 121 SSGQNTDEAQTOSTISRFACRIYCDRNEPYTARIFAAGFDSKNIPLGEKAKMKNPDG 180
QY 181 HMDGLTNGVLVMPHPRGFTBESQPGVWREISVCGDYTLRRTSAQORGLVSESTNVL 240
DB 181 HMDGLTNGVLVMPHPRGFTBESQPGVWREISVCGDYTLRRTSAQORGLVSESTNVL 240
QY 241 QDGSLLDLCGATLLMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
DB 241 QDGSLLDLCGATLLMTADGLFHTPTQKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
QY 301 VVEKQPMAYLSCGHVGHYNNWGHRSOTENARECPMCRITVGPVPLMTGCEAGFYVDAG 360
DB 301 VVEKQPMAYLSCGHVGHYNNWGHRSOTENARECPMCRITVGPVPLMTGCEAGFYVDAG 360
QY 361 PPTHAFPCGHVSEKSAKYWSQIPLPHGTHFAHACPCATQLVGEONCIKLIPOGPID 420
DB 361 PPTHAFPCGHVSEKSAKYWSQIPLPHGTHFAHACPCATQLVGEONCIKLIPOGPID 420

RESULT 3

ABR41083
ID ABR41083 standard; protein; 420 AA.

XX
AC ABR41083;

DT 22-MAY-2003 (first entry)

XX
DE Human MAP kinase cascade activator #59.

XX
KW Human; Elki phosphorylation; Elki phosphorylation kinase; virucide;
KW antiinflammatory; immunomodulator; cytotoxic; antiallergic; anti-HIV;
KW antineumatic; antiarthritic; antidiabetic; antiaesthetic; gene therapy;
KW inflammation; autoimmune disease; viral disease; cancer; diabetes;
KW rheumatoid arthritis; asthma; allergic rhinitis; AIDS; viral hepatitis;
KW IGA nephritis.

XX
OS Homo sapiens.

XX MO2003008589-A1.
 XX 30-JAN-2003.
 XX 15-JUL-2002; 2002MO-JP007174.
 XX 18-JUL-2001; 2001JP-00218204.
 XX 31-AUG-2001; 2001JP-00263450.
 XX 21-JAN-2002; 2002JP-00012176.
 XX (ASAH) ASAMI KASEI KOGYO KK.
 XX Matsuzaki O, Matsuda A, Nagano Y, Suzuki N;
 XX WPI; 2003-229582/22.
 XX N-PSDB; ACC42349.
 XX Etk1 phosphorylation-associated gene and its encoded protein with MAP
 PT kinase cascade effect, applicable in diagnosis of and developing drugs
 PT for e.g. inflammation, autoimmune diseases, viral diseases and cancer.
 XX
 XX Claim 1; Page 640-643; 762pp; Japanese.
 XX
 XX The invention relates to a novel purified protein having Etk1
 CC phosphorylation activity and/or an activity of activating Etk1
 CC phosphotyrosine kinase. A protein of the invention has antiinflammatory,
 CC immunomodulator, virucide, cytostatic, antiallergic, antineoplastic,
 CC antitumor, antidiabetic, antiasomatic, and anti-HIV activity. The
 CC polynucleotides may have a use in gene therapy. The gene and its encoded
 CC protein are applicable in diagnosis of and developing drugs for e.g.
 CC inflammation, autoimmune diseases, viral diseases and cancer such as
 CC rheumatoid arthritis, diabetes, asthma, allergic rhinitis, AIDS, viral
 CC hepatitis and Iga nephritis. The present sequence is used in the
 CC exemplification of the invention.
 CC
 XX Sequence 420 AA:
 SQ
 Query Match 100.0%; Score 2290; DB 6; Length 420;
 Best Local Similarity 100.0%; Pred. No. 2.8e-226;
 Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
 QY 1 MFSPGQSEHCAPNKEPVKGEIVLVGYNGALPNDGRGRKRSFALYRPRKANGVKESTVH 60
 DB 1 MFSPGQSEHCAPNKEPVKGEIVLVGYNGALPNDGRGRKRSFALYRPRKANGVKESTVH 60
 QY 61 VISTPQASKAISCKGQHSISYTLSSRNQTVVVEYTHDQTMFQVGRSTESPIDVVTDTI 120
 DB 61 VISTPQASKAISCKGQHSISYTLSSRNQTVVVEYTHDQTMFQVGRSTESPIDVVTDTI 120
 QY 121 SSGQNTDEAQTOSTISRPAFCRIYCDNEPYTARIFAAGPDSKNIFLGEKRAAKKNDG 180
 DB 121 SSGQNTDEAQTOSTISRPAFCRIYCDNEPYTARIFAAGPDSKNIFLGEKRAAKKNDG 180
 QY 181 HMOGLTNGVLVWHPRGFTFESQPVWREISVCGDVYTLRETSSAQRGLVSESTNVL 240
 DB 181 HMOGLTNGVLVWHPRGFTFESQPVWREISVCGDVYTLRETSSAQRGLVSESTNVL 240
 QY 241 QDSGLIDLCATLLMRTADGLFHTPTOKHTEALROEINARPOCPVGLNTLAFPSINRKE 300
 DB 241 QDSGLIDLCATLLMRTADGLFHTPTOKHTEALROEINARPOCPVGLNTLAFPSINRKE 300
 QY 301 VVEKQPMAYLSCGHVGHYNWGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
 DB 301 VVEKQPMAYLSCGHVGHYNWGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
 QY 361 PPTHTAFTPCGHVCEKSAKXWSQIPLPHGTHTAFAACPFATOLVGBONCIKLIPOPID 420
 DB 361 PPTHTAFTPCGHVCEKSAKXWSQIPLPHGTHTAFAACPFATOLVGBONCIKLIPOPID 420

ID ADC31180 standard; protein; 420 AA.
 XX
 AC ADC31180;
 DT 16-DEC-2003 (first entry)
 XX
 DE Human novel polypeptide sequence, SEQ ID NO:1262.
 XX
 KW Human, diagnostic; drug screening; forensics; gene mapping;
 KW biodiversity assessment; Parkinson's disease; Alzheimer's disease;
 KW neurodegenerative diseases; anaemia; platelet disorder; wound; burns;
 KW ulcers; osteoporosis; autoimmune disease; cancer;
 KW molecular weight marker; food supplement; antiparkinsonian; nootropic;
 KW neuroprotective; anti-nausea; anticoagulant; thrombolytic; vulnerary;
 KW antidiabetic; osteopathic; immunosuppressive; antiinflammatory; cytostatic;
 KW gene therapy; chromosome 14q21.
 XX
 OS Homo sapiens.
 XX
 PN MO2003029271-A2.
 XX
 PD 10-APR-2003.
 XX
 PF 24-SEP-2002; 2002MO-US030474.
 XX
 PR 24-SEP-2001; 2001US-0324631P.
 XX
 PA (HYSEQ-) HYSEQ INC.
 XX
 PI Tang TY, Zhang J, Ren F, Xue AJ, Zhao QA, Wang J, Wehrman T;
 PI Zhou P, Ghosh M, Wang D, Ma Y, Asundi V, Wang Z, Weng G;
 PI Haley-Vicente D, Dymanc RT;
 XX
 DR WPI; 2003-371981/35.
 DR N-PSDB; ADC30209.
 XX
 PT New polynucleotide and polypeptide useful for diagnosing, preventing or
 PT treating conditions such as neurodegenerative diseases, anemias, platelet
 PT disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 PT cancer.
 XX
 PS Claim 20; SEQ ID NO 1262; 1185pp; English.
 XX
 CC The invention relates to 971 novel human cDNA sequences (ADC29919-
 CC ADC30889) and the polypeptides they encode (ADC30890-ADC31860). The
 CC invention also relates to nucleic acid sequences over 99% identical with
 CC the novel human cDNAs. The invention additionally encompasses expression
 CC vectors and host cells comprising a nucleic acid of the invention; the
 CC recombinant production of a polypeptide of the invention; an antibody
 CC against a polypeptide of the invention; a method of detecting
 CC polynucleotides or polypeptides of the invention; and methods of
 CC identifying a compound which binds to a polypeptide of the invention. The
 CC invention further discloses methods of preventing, treating or
 CC ameliorating a medical condition; kits comprising polynucleotide probes
 CC and/or monoclonal antibodies for carrying out the methods of the
 CC invention; methods for the identification of compounds that modulate the
 CC expression or activity of the polynucleotide and/or polypeptide; and 767
 CC contig sequences corresponding to the cDNA sequences of the invention
 CC (ADC31861-ADC32627) and the polypeptides encoded by the contigs (ADC32628
 CC -ADC33394). The nucleic acids and polypeptides of the invention are
 CC useful in diagnostics, drug screening, forensics, gene mapping, in the
 CC identification of mutations responsible for genetic disorders or other
 CC traits, for assessing biodiversity, and in producing many other types of
 CC data and products dependent on DNA and amino acid sequences. They are
 CC also used for treating diseases such as Parkinson's disease, Alzheimer's
 CC disease and other neurodegenerative diseases, anaemia, platelet
 CC disorders, wounds, burns, ulcers, osteoporosis, autoimmune diseases or
 CC cancer. The nucleic acids may also be used as hybridisation probes or
 CC primers, and in the recombinant production of a protein. The polypeptides
 CC are also useful in generating antibodies, as molecular weight markers,
 CC and as food supplements. The present sequence represents a specifically
 CC claimed human polypeptide sequence of the invention. Note: The sequence
 CC data for this patent did not form part of the printed specification, but

CC was obtained in electronic format directly from WIPO at
 CC ftp.wipo.int/pub/published_pct_sequences.

XX Sequence 420 AA;

Query Match 100.0%; Score 2290; DB 7; Length 420;
 Best Local Similarity 100.0%; Pred. No. 2.8e-226;
 Matches 420; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MESPGEHCAPNKEPVKYGELVLTGYNGALPNGDRGRKSRFPALYKRPANGVKPSTVH 60
 DB 1 MESPGEHCAPNKEPVKYGELVLTGYNGALPNGDRGRKSRFPALYKRPANGVKPSTVH 60
 QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVVEYTHDKDMDPQVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQASKAISCKGHSISYTLSSRNQTVVVEYTHDKDMDPQVGRSTESPIDFVVTDTI 120
 QY 121 SSGQNTDEAQITOSTISRACRIVCDBRNEPYTARIIPAAGFDSKNIPLGEKAKMKNPDG 180
 DB 121 SSGQNTDEAQITOSTISRACRIVCDBRNEPYTARIIPAAGFDSKNIPLGEKAKMKNPDG 180
 QY 181 HMDGLTTNGVLVNHPRGFTBESQPGVWREISVCGDVYTLRETRSAQQRKLVESSTNVL 240
 DB 181 HMDGLTTNGVLVNHPRGFTBESQPGVWREISVCGDVYTLRETRSAQQRKLVESSTNVL 240
 QY 241 QDGSLLIDLCGATLMTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
 DB 241 QDGSLLIDLCGATLMTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
 QY 301 VVEKQPMAYLSCGHVGHYNNWGHRSDEANERECPCRTVGYPVPLMLGCEAGFYVDAG 360
 DB 301 VVEKQPMAYLSCGHVGHYNNWGHRSDEANERECPCRTVGYPVPLMLGCEAGFYVDAG 360
 QY 361 PPTHAFPCGHVCEKSAKYWSQIPLPHGTHAFHAACPFCATQVGEONCIKILIFQGPID 420
 DB 361 PPTHAFPCGHVCEKSAKYWSQIPLPHGTHAFHAACPFCATQVGEONCIKILIFQGPID 420

RESULT 5

ABB07922 standard; protein; 420 AA.

AC ABB07922;

DT 30-JUL-2002 (first entry)

DE Human pellino-2 polypeptide.

XX Nuclear factor-kappaB; NF-kappaB; p38; pellino; interleukin-1; vitruclide;
 KM antibacterial; fungicide; protozoacidal; antiaesthetic; antirheumatic;
 KM antiarthritic; antiinflammatory; antiatherosclerotic; neuroprotective;
 KM nootropic; antidiabetic; human; pellino-2.

OS Homo sapiens.

XX MO200183739-A2.

PN 08-NOV-2001.

XX 27-APR-2001; 2001WO-US013676.

XX 28-APR-2000; 2000US-0200198P.

XX (IMMV) IMMUNEX CORP.

XX Bird TA, Cosman DJ;

XX WPI; 2002-066532/09.

XX N-PSDB; ABL58449.

PT New Pellino polypeptides for identifying compounds that alter polypeptide
 PT activity, treating pathogenic infection or inhibiting apoptosis, are
 PT capable of stimulating nuclear factor-kappaB- or p38-dependent

PT tranacscription.

XX Claim 1; Page 61-63; 70pp; English.

CC The invention provides polypeptides capable of stimulating nuclear factor
 CC (NF)-kappaB-dependent transcription or p38-dependent transcription,
 CC referred as Pellino polypeptides. The pellino polypeptides are useful for
 CC identifying modulators that alter the pellino polypeptide and pellino
 CC dominant-negative activity. They are also useful for identifying
 CC compounds that inhibit the binding activity of the polypeptides and to
 CC study cell-signal transduction. They are useful for preventing or
 CC treating infection by a pathogen such as virus, bacterial, fungi, algae
 CC or protozoa, or inhibiting apoptosis. Dominant-negative pellino
 CC polypeptides are useful for treating inflammatory conditions such as
 CC asthma, rheumatoid arthritis, inflammatory bowel disease, Crohn's
 CC disease, ulcerative colitis, atherosclerosis and Alzheimer's disease, and
 CC also for inhibiting mitogen activated protein (MAP) kinase-activated
 CC pathways. Pellino polypeptides and polynucleotides are useful to identify
 CC small molecule inhibitors of protein association or function of Pellino,
 CC and other molecules involved in interleukin (IL)-1 signaling. The present
 CC sequence represents a human pellino-2 polypeptide

XX Sequence 420 AA;

Query Match 99.7%; Score 2283; DB 5; Length 420;
 Best Local Similarity 99.8%; Pred. No. 1.5e-225;
 Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MESPGEHCAPNKEPVKYGELVLTGYNGALPNGDRGRKSRFPALYKRPANGVKPSTVH 60
 DB 1 MESPGEHCAPNKEPVKYGELVLTGYNGALPNGDRGRKSRFPALYKRPANGVKPSTVH 60
 QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVVEYTHDKDMDPQVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQASKAISCKGHSISYTLSSRNQTVVVEYTHDKDMDPQVGRSTESPIDFVVTDTI 120
 QY 121 SSGQNTDEAQITOSTISRACRIVCDBRNEPYTARIIPAAGFDSKNIPLGEKAKMKNPDG 180
 DB 121 SSGQNTDEAQITOSTISRACRIVCDBRNEPYTARIIPAAGFDSKNIPLGEKAKMKNPDG 180
 QY 181 HMDGLTTNGVLVNHPRGFTBESQPGVWREISVCGDVYTLRETRSAQQRKLVESSTNVL 240
 DB 181 HMDGLTTNGVLVNHPRGFTBESQPGVWREISVCGDVYTLRETRSAQQRKLVESSTNVL 240
 QY 241 QDGSLLIDLCGATLMTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
 DB 241 QDGSLLIDLCGATLMTADGLFHTPTOKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
 QY 301 VVEKQPMAYLSCGHVGHYNNWGHRSDEANERECPCRTVGYPVPLMLGCEAGFYVDAG 360
 DB 301 VVEKQPMAYLSCGHVGHYNNWGHRSDEANERECPCRTVGYPVPLMLGCEAGFYVDAG 360
 QY 361 PPTHAFPCGHVCEKSAKYWSQIPLPHGTHAFHAACPFCATQVGEONCIKILIFQGPID 420
 DB 361 PPTHAFPCGHVCEKSAKYWSQIPLPHGTHAFHAACPFCATQVGEONCIKILIFQGPID 420

RESULT 6

ADP48671 standard; protein; 420 AA.

AC ADP48671;

DT 09-SEP-2004 (first entry)

XX Human Pellino-2 protein SEQ ID NO:8.

XX nuclear factor kappa B dependent transcription inhibitor;
 KM NF-kB-dependent transcription inhibitor;
 KM p38-dependent transcription inhibitor; Pellino-1;
 KM interleukin 1 receptor-associated kinase 4;
 KM IL-1 receptor-associated kinase 4; IRAK-4; antiarteriosclerotic;
 KM antiarthritic; antiaesthetic; antiinflammatory; antidiabetic; antidiabetic;

KM Gastrointestinal; neuroprotective; nootropic; IL-1 antagonist;
KM IRAK protein kinase family inhibitor; asthma; rheumatoid arthritis;
KM NF kappa B inducing kinase inhibitor; Crohn's disease; ulcerative colitis;
KM inflammatory bowel disease; Crohn's disease; ulcerative colitis;
KM atherosclerosis; Alzheimer's disease; human; Pellino-2; chromosome 14.
OS Homo sapiens.
PN MO2004053092-A2.
PD 24-JUN-2004.
PF 09-DEC-2003; 2003WO-US039188.
PR 11-DEC-2002; 2002US-00317250.
XX (IMMUNEX CORP.
PA (CLEVELAND CLINIC FOUND.
PI Bird TA, Cosman DJ, Li X;
XX WPI; 2004-460927/45.
DR N-PSDB; ADP48670.
XX
PT Identifying inhibitors of IL-1 signaling, useful for treating e.g.,
PT asthma, and rheumatoid arthritis, comprises assaying the association of a
PT Pellino-1 polypeptide with IRAK-4 in the presence of a test compound.
PS Example 1; SEQ ID NO 8; 81pp; English.
XX
XX The present invention describes a method for identifying compounds (C)
CC that inhibit nuclear factor kappa B (NF-kB)-dependent transcription or
CC p38-dependent transcription. The method comprises: (1) mixing a test
CC compound with a Pellino-1 polypeptide; (b) assaying the association of
CC the Pellino-1 polypeptide with IL-1 receptor-associated kinase 4 (IRAK-4)
CC in the presence of the test compound; and (c) determining whether the
CC test compound inhibits the association of the Pellino-1 polypeptide with
CC a binding partner. Also described: (1) an inhibitory nucleic acid that
CC binds to a nucleic acid encoding an amino acid sequence as described
CC above, where the presence of the inhibitory nucleic acid within a cell
CC inhibits the association of Pellino-1 with IRAK-4; and (2) an inhibitory
CC polypeptide comprising an antibody fragment that binds to a polypeptide
CC comprising an amino acid sequence as described above, where the presence
CC of the inhibitory polypeptide within a cell inhibits the association of
CC Pellino-1 with IRAK-4. (C) have antiarteriosclerotic, antiarthritic,
CC gastroenteric, neuroprotective, antiinflammatory, antitumor, anticancer,
CC used as an IL-1 antagonist, IRAK protein kinase family inhibitor, and NF
CC kappa B inducing kinase inhibitor. The methods and compositions of the
CC present invention are useful for the prevention and/or treatment of
CC diseases or conditions associated with aberrant expression or activity of
CC the Pellino-1 polypeptide, such as asthma, rheumatoid arthritis,
CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
CC atherosclerosis and Alzheimer's disease. The present sequence represents
CC human Pellino-2, which is used in the exemplification of the present
CC invention. The human Pellino-2 gene is located on chromosome 14, more
CC specifically to 14q24.3.
XX
XX Sequence 420 AA:
SQ
Query Match 99.7%; Score 2283; DB 8; Length 420;
Best Local Similarity 99.8%; Pred. No. 1.5e-225;
Matches 419; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MFSPGQEHGCAPNKEPVKYGELVLTGYNGALPNDGRGRSRFPALYKRPKANGKPESTVH 60
Db 1 MFSPGQEHGCAPNKEPVKYGELVLTGYNGALPNDGRGRSRFPALYKRPKANGKPESTVH 60
QY 61 VISTPQASKAISCKGHSISYTLSSRNQTVVVEYTHDQTDMPFGVGRSTESPDPVVDTTI 120
Db 61 VISTPQASKAISCKGHSISYTLSSRNQTVVVEYTHDQTDMPFGVGRSTESPDPVVDTTI 120
QY 121 SSGQNTDEAQITOSTISRPRACIVCDRNEPYTARIPAAFGDSSKNITLGEKAAKKNPDG 180
Db 121 SSGQNTDEAQITOSTISRPRACIVCDRNEPYTARIPAAFGDSSKNITLGEKAAKKNPDG 180

Db 121 SSGQNTDEAQITOSTISRPRACIVCDRNEPYTARIPAAFGDSSKNITLGEKAAKKNPDG 180
QY 181 HMDGLTTNGVLVNHPRGFTTESQOPVMEISVGVGYTLRETRSAQQRKLVESSTNVL 240
Db 181 HMDGLTTNGVLVNHPRGFTTESQOPVMEISVGVGYTLRETRSAQQRKLVESSTNVL 240
QY 241 QDGLSLDLCGATLLMTAQLFTTPQKHIALROBINARPCCPVGLMTLAPPSINRKE 300
Db 241 QDGLSLDLCGATLLMTAQLFTTPQKHIALROBINARPCCPVGLMTLAPPSINRKE 300
QY 301 VVEKQPMAYLSCGHVGYNMGHRSPTANERECMCRTVGYPVLMIGCEAGFYVDAG 360
Db 301 VVEKQPMAYLSCGHVGYNMGHRSPTANERECMCRTVGYPVLMIGCEAGFYVDAG 360
QY 361 PPTHAFTPCGHVCEKSAKXYSQIPPHGTHAFHAACPFCATQLVGBQNCIKLIPOGPID 420
Db 361 PPTHAFTPCGHVCEKSAKXYSQIPPHGTHAFHAACPFCATQLVGBQNCIKLIPOGPID 420
RESULT 7
AB07921
ID ABB07921 standard; protein; 419 AA.
XX
AC ABB07921;
XX
DT 30-JUL-2002 (first entry)
XX
XX Murine pellino-2 polypeptide.
XX
KM Nuclear factor-kappaB; NF-kappaB; p38; pellino; interleukin-1; virucide;
KM antibacterial; fungicide; protozoacide; antiasthmatic; antirheumatic;
KM antiarthritic; antiinflammatory; antiatherosclerotic; neuroprotective;
KM nootropic; antileuk; mouse; pellino-2.
XX
OS Mus musculus.
XX
PN MO200183739-A2.
XX
PD 08-NOV-2001.
XX
PF 27-APR-2001; 2001MO-US013676.
XX
XX 28-APR-2000; 2000US-0200198P.
XX
XX (IMMUNEX CORP.
PA (IMMUNEX CORP.
PI Bird TA, Cosman DJ;
XX WPI; 2002-066532/09.
DR N-PSDB; ABL58448.
XX
XX New Pellino polypeptides for identifying compounds that alter polypeptide
PT activity, treating pathogenic infection or inhibiting apoptosis, are
PT capable of stimulating nuclear factor-kappaB- or p38-dependent
PT transcription.
XX
XX Example; Page 59-60; 70pp; English.
XX
XX The invention provides polypeptides capable of stimulating nuclear factor
CC (NF)-kappaB-dependent transcription or p38-dependent transcription,
CC referred as Pellino polypeptides. The Pellino polypeptides are useful for
CC identifying modulators that alter the Pellino polypeptide and Pellino
CC dominant-negative activity. They are also useful for identifying
CC compounds that inhibit the binding activity of the polypeptides and to
CC study cell-signal transduction. They are useful for preventing or
CC treating infection by a pathogen such as virus, bacterial, fungi, algae
CC or protozoa, or inhibiting apoptosis. Dominant-negative Pellino
CC polypeptides are useful for treating inflammatory conditions such as
CC asthma, rheumatoid arthritis, inflammatory bowel disease, Crohn's
CC disease, ulcerative colitis, atherosclerosis and Alzheimer's disease, and
CC also for inhibiting mitogen activated protein (MAP) kinase-activated
CC pathways. Pellino polypeptides and polynucleotides are useful to identify

CC small molecule inhibitors of protein association or function of Pellino,
 CC and other molecules involved in interleukin (IL)-1 signaling. The present
 CC sequence represents a murine pellino-2 polypeptide

SO Sequence 419 AA;

Query Match 95.0%; Score 2175.5; DB 5; Length 419;
 Best Local Similarity 95.0%; Pred. No. 1,7e-214;
 Matches 399; Conservative 10; Mismatches 10; Indels 1; Gaps 1;

QY 1 MESPQGEHCAPKPEVKYGEIVLVGYNGALPVGDRGRKSRPALYKRKANGKSTYH 60
 DB 1 MESPQGEHCAPKPEVKYGEIVLVGYNGALPVGDRGRKSRPALYKRKANGKSTYH 60
 QY 61 VISTPQAKAISCKGHSISYTLNRNQTVEYTHDKDTMPOVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQAKAISCKGHSISYTLNRNQTVEYTHDKDTMPOVGRSTESPIDFVVTDTI 120
 QY 61 MVTTPQAKAISCKGHSISYTLNRNQTVEYTHDKDTMPOVGRSTESPIDFVVTDTI 120
 DB 61 MVTTPQAKAISCKGHSISYTLNRNQTVEYTHDKDTMPOVGRSTESPIDFVVTDTI 120
 QY 121 SSGQNTDEAQTOSTISRFACRIVCDBNEPYTARIFAAGFSSKNIFLGEKAKWKXNDG 180
 DB 121 SSGQNTDEAQTOSTISRFACRIVCDBNEPYTARIFAAGFSSKNIFLGEKAKWKXNDG 180
 QY 181 HMGGLTNGVLVWHPGCGFTESQPGVWREISVCGDYTTRETRSAQGRKLYSESTNVL 240
 DB 181 HMGGLTNGVLVWHPGCGFTESQPGVWREISVCGDYTTRETRSAQGRKLYSESTNVL 240
 QY 241 QDGLSLIDLCATLMTADGLFTPTQKIEALROEINARPOCPVGLNTLAFPSINRKE 300
 DB 241 QDGLSLIDLCATLMTADGLFTPTQKIEALROEINARPOCPVGLNTLAFPSINRKE 300
 QY 301 VVEKQPMAYLSCGHVGYNMGHRSDEANERECPCRTVGYPVPLMGCENGFYVDAG 360
 DB 301 VVEKQPMAYLSCGHVGYNMGHRSDEANERECPCRTVGYPVPLMGCENGFYVDAG 360
 QY 300 VVEKQPMAYLSCGHVGYNMGHRSDEANERECPCRTVGYPVPLMGCENGFYVDAG 359
 DB 300 VVEKQPMAYLSCGHVGYNMGHRSDEANERECPCRTVGYPVPLMGCENGFYVDAG 359
 QY 361 PPTHTAPCGHVCSEKSAKYSQIPLPHGTHAFHAACPFCATQVGEONCICKLIFQGPID 420
 DB 361 PPTHTAPCGHVCSEKSAKYSQIPLPHGTHAFHAACPFCATQVGEONCICKLIFQGPID 420
 QY 360 PPTHTAPCGHVCSEKSAKYSQIPLPHGTHAFHAACPFCATQVGEONCICKLIFQGPID 419
 DB 360 PPTHTAPCGHVCSEKSAKYSQIPLPHGTHAFHAACPFCATQVGEONCICKLIFQGPID 419

RESULT 8

ID ADP48669 standard; protein; 419 AA.

AC ADP48669;

DT 09-SEP-2004 (first entry)

DE Mouse Pellino-2 protein SEQ ID NO:6.

XX nuclear factor kappa B dependent transcription inhibitor;
 XX NF-kB-dependent transcription inhibitor;
 XX p38-dependent transcription inhibitor; Pellino-1;
 XX interleukin 1 receptor-associated kinase 4;
 XX IL-1 receptor-associated kinase 4; IRAK-4; antiarteriosclerotic;
 XX antiarthritic; antiasthmatic; antiinflammatory; antirheumatic;
 XX gastrointestinal; neuroprotective; nocotropic; IL-1 antagonist;
 XX IRAK protein kinase family inhibitor;
 XX NF kappa B inducing kinase inhibitor; asthma; rheumatoid arthritis;
 XX inflammatory bowel disease; Crohn's disease; ulcerative colitis;
 XX atherosclerosis; Alzheimer's disease; mouse; Pellino-2.

OS Mus musculus.

PN W02004053092-A2.

PD 24-JUN-2004.

PF 09-DEC-2003; 2003MO-US039188.

PR 11-DEC-2002; 2002US-00317250.

XX (IMMUNEX CORP.

PA (CLEVELAND CLINIC FOUND.

XX Bird TA, Cosman DJ, Li X;
 PI WPI; 2004-480927/45.
 DR N-PSDB; ADP48668.

XX Identifying inhibitors of IL-1 signaling, useful for treating e.g.,
 PT asthma, and rheumatoid arthritis, comprises assaying the association of a
 PT Pellino-1 polypeptide with IRAK-4 in the presence of a test compound.
 XX Example 1; SEQ ID NO 6; 81pp; English.

PS The present invention describes a method for identifying compounds (C)
 XX that inhibit nuclear factor kappa B (NF-kB)-dependent transcription or
 CC p38-dependent transcription. The method comprises: (1) mixing a test
 CC compound with a Pellino-1 polypeptide; (b) assaying the association of
 CC the Pellino-1 polypeptide with IL-1 receptor-associated kinase 4 (IRAK-4)
 CC in the presence of the test compound; and (c) determining whether the
 CC test compound inhibits the association of the Pellino-1 polypeptide with
 CC a binding partner. Also described: (1) an inhibitory nucleic acid that
 CC binds to a nucleic acid encoding an amino acid sequence as described
 CC above, where the presence of the inhibitory nucleic acid within a cell
 CC inhibits the association of Pellino-1 with IRAK-4; and (2) an inhibitory
 CC polypeptide comprising an antibody fragment that binds to a polypeptide
 CC comprising an amino acid sequence as described above, where the presence
 CC of the inhibitory polypeptide within a cell inhibits the association of
 CC Pellino-1 with IRAK-4. (C) have antiarteriosclerotic, antiarthritic,
 CC antiasthmatic, antiinflammatory, antirheumatic, antitumor,
 CC gastrointestinal, neuroprotective and nocotropic activities, and can be
 CC used as an IL-1 antagonist, IRAK protein kinase family inhibitor, and NF
 CC kappa B inducing kinase inhibitor. The methods and compositions of the
 CC present invention are useful for the prevention and/or treatment of
 CC diseases or conditions associated with aberrant expression or activity of
 CC the Pellino-1 polypeptide, such as asthma, rheumatoid arthritis,
 CC inflammatory bowel disease, Crohn's disease, ulcerative colitis,
 CC atherosclerosis and Alzheimer's disease. The present sequence represents
 CC mouse Pellino-2, which is used in the exemplification of the present
 CC invention.

SO Sequence 419 AA;

Query Match 95.0%; Score 2175.5; DB 8; Length 419;
 Best Local Similarity 95.0%; Pred. No. 1,7e-214;
 Matches 399; Conservative 10; Mismatches 10; Indels 1; Gaps 1;

QY 1 MESPQGEHCAPKPEVKYGEIVLVGYNGALPVGDRGRKSRPALYKRKANGKSTYH 60
 DB 1 MESPQGEHCAPKPEVKYGEIVLVGYNGALPVGDRGRKSRPALYKRKANGKSTYH 60
 QY 61 VISTPQAKAISCKGHSISYTLNRNQTVEYTHDKDTMPOVGRSTESPIDFVVTDTI 120
 DB 61 VISTPQAKAISCKGHSISYTLNRNQTVEYTHDKDTMPOVGRSTESPIDFVVTDTI 120
 QY 61 MVTTPQAKAISCKGHSISYTLNRNQTVEYTHDKDTMPOVGRSTESPIDFVVTDTI 120
 DB 61 MVTTPQAKAISCKGHSISYTLNRNQTVEYTHDKDTMPOVGRSTESPIDFVVTDTI 120
 QY 121 SSGQNTDEAQTOSTISRFACRIVCDBNEPYTARIFAAGFSSKNIFLGEKAKWKXNDG 180
 DB 121 SSGQNTDEAQTOSTISRFACRIVCDBNEPYTARIFAAGFSSKNIFLGEKAKWKXNDG 180
 QY 181 HMGGLTNGVLVWHPGCGFTESQPGVWREISVCGDYTTRETRSAQGRKLYSESTNVL 240
 DB 181 HMGGLTNGVLVWHPGCGFTESQPGVWREISVCGDYTTRETRSAQGRKLYSESTNVL 240
 QY 241 QDGLSLIDLCATLMTADGLFTPTQKIEALROEINARPOCPVGLNTLAFPSINRKE 300
 DB 241 QDGLSLIDLCATLMTADGLFTPTQKIEALROEINARPOCPVGLNTLAFPSINRKE 300
 QY 301 VVEKQPMAYLSCGHVGYNMGHRSDEANERECPCRTVGYPVPLMGCENGFYVDAG 360
 DB 301 VVEKQPMAYLSCGHVGYNMGHRSDEANERECPCRTVGYPVPLMGCENGFYVDAG 360
 QY 300 VVEKQPMAYLSCGHVGYNMGHRSDEANERECPCRTVGYPVPLMGCENGFYVDAG 359
 DB 300 VVEKQPMAYLSCGHVGYNMGHRSDEANERECPCRTVGYPVPLMGCENGFYVDAG 359
 QY 361 PPTHTAPCGHVCSEKSAKYSQIPLPHGTHAFHAACPFCATQVGEONCICKLIFQGPID 420
 DB 361 PPTHTAPCGHVCSEKSAKYSQIPLPHGTHAFHAACPFCATQVGEONCICKLIFQGPID 420
 QY 360 PPTHTAPCGHVCSEKSAKYSQIPLPHGTHAFHAACPFCATQVGEONCICKLIFQGPID 419
 DB 360 PPTHTAPCGHVCSEKSAKYSQIPLPHGTHAFHAACPFCATQVGEONCICKLIFQGPID 419

RESULT 9
 ABR41082
 ID ABR41082 standard; protein; 419 AA.
 XX
 AC ABR41082;
 XX
 DT 22-MAY-2003 (first entry)
 XX
 DE Mouse MAP kinase cascade activator #9.
 XX
 KW Mouse; Etk1 phosphorylation; Etk1 phosphorylation kinase; virucide;
 KW antiinflammatory; immunomodulator; cytostatic; antiallergic; anti-HIV;
 KW antineumatic; antiarthritic; antidiabetic; antiaesthetic; gene therapy;
 KW inflammation; autoimmune disease; viral disease; cancer; diabetes;
 KW rheumatoid arthritis; asthma; allergic rhinitis; AIDS; viral hepatitis;
 KW Iga nephritis.
 XX
 OS Mus musculus.
 XX
 PN WO2003008589-A1.
 XX
 PD 30-JAN-2003.
 XX
 PF 15-JUL-2002; 2002WO-JP007174.
 XX
 PR 18-JUL-2001; 2001JP-00218204.
 PR 31-AUG-2001; 2001JP-00263450.
 PR 21-JAN-2002; 2002JP-00012176.
 XX
 PA (ASAH) ASAH KASEI KOGYO KK.
 XX
 PI Matsuzaki O, Matsuda A, Nagano Y, Suzuki N;
 XX
 DR WPI; 2003-229582/22.
 DR N-PSDB; ACC42348.
 XX
 PT Etk1 phosphorylation-associated gene and its encoded protein with MAP
 PT kinase cascade effect, applicable in diagnosis of and developing drugs
 PT for e.g. inflammation, autoimmune disease, viral diseases and cancer.
 XX
 PS Claim 1; Page 628-630; 762pp; Japanese.
 XX
 CC The invention relates to a novel purified protein having Etk1
 CC phosphorylation activity and/or an activity of activating Etk1
 CC phosphorylation kinase. A protein of the invention has antiinflammatory,
 CC immunomodulator, virucide, cytostatic, antiallergic, antineumatic,
 CC antiarthritic, antidiabetic, antiaesthetic, and anti-HIV activity. The
 CC polynucleotides may have a use in gene therapy. The gene and its encoded
 CC protein are applicable in diagnosis of and developing drugs for e.g.
 CC inflammation, autoimmune diseases, viral diseases and cancer such as
 CC rheumatoid arthritis, diabetes, asthma, allergic rhinitis, AIDS, viral
 CC hepatitis and Iga nephritis. The present sequence is used in the
 CC exemplification of the invention
 XX
 SQ Sequence 419 AA.
 Query Match 92.4%; Score 2111.5; DB 6; Length 419;
 Best Local Similarity 92.4%; Pred. No. 6.6e-208;
 Matches 388; Conservative 12; Mismatches 19; Indels 1; Gaps 1;
 QY 1 MESPGEHCAPEKPEYKELVYLVGNGALPNDGRKRRKSPALYRPRANGVPESTVH 60
 DB 1 MESPGEHCAPEKPEYKELVYLVGNGALPNDGRKRRKSPALYRPRANGVPESTVH 60
 QY 61 VISTPQSKAISCQHSISYTLNRCTVVEYTHDQTMFOVGRSTESPIDPVVDTI 120
 DB 61 MVTTPQSKAISCQHSISYTLNRCTVVEYTHDQTMFOVGRSTESPIDPVVDTI 120
 QY 121 SGGQNTDEAQITOSTISRFACRIVCDNEPYTARIFPAGFDSSKNITLGEKAAKKNPDG 180
 DB 121 SGGQNTDEAQITOSTISRFACRIVCDNEPYTARIFPAGFDSSKNITLGEKAAKKNPDG 179

QY 181 HMDGLTNGVLVNHPRGFTSESPGCVWREISVCGDVYTLRETRSAQORGLVSESTNVL 240
 DB 180 HMDGLTNGVLVNHPRGFTSESPGCVWREISVCGDVYTLRETRSAQORGLVSESTNVL 239
 QY 241 QDGLIDLCATLIMRTADGLFTPTQKHIEALROEINARPOCPVGLNTLAPPSINRKE 300
 DB 240 QDGLIDLCATLIMRTADGLFTPTQKHIEALROEINARPOCPVGLNTLAPPSINRKE 299
 QY 301 VVEKOPWATLSCGHVGYNNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 360
 DB 300 VVEKOPWATLSCGHVGYNNMGRSDTEANERECPCRTVGPVPLMLGCEAGFYVDAG 359
 QY 361 PPTHAFFPCGVCSKSAKYSQIPLPHGTHAFAACPFCATQLVGRQNCIKLIFQSPID 420
 DB 360 XXXHVFPCGVCSKSAKYSQIPLPHGTHAFAACPFCATQLVGRQNCIKLIFQSPID 419
 RESULT 10
 ABB07920
 ID ABB07920 standard; protein; 418 AA.
 XX
 AC ABB07920;
 XX
 DT 30-JUL-2002 (first entry)
 XX
 DE Human pellino-1 polypeptide.
 XX
 KW Nuclear factor-kappaB; NF-kappaB; p38; pellino; interleukin-1; virucide;
 KW antibacterial; fungicide; fungicide; protozoacide; antisthmatic; antirheumatic;
 KW antiarthritic; antineumatic; antiaesthetic; antiatherosclerotic; neuroprotective;
 KW nootropic; antileuc; human; pellino-1.
 XX
 OS Homo sapiens.
 XX
 PN WO200183739-A2.
 XX
 PD 08-NOV-2001.
 XX
 PF 27-APR-2001; 2001WO-US013676.
 XX
 PR 28-APR-2000; 2000US-0200198P.
 XX
 PA (IMMUNEX CORP.
 XX
 PI Bird TA, Cosman DJ;
 XX
 DR WPI; 2002-066532/09.
 DR N-PSDB; ABL58447.
 XX
 PT New pellino polypeptides for identifying compounds that alter polypeptide
 PT activity, treating pathogenic infection or inhibiting apoptosis, are
 PT capable of stimulating nuclear factor-kappaB- or p38-dependent
 PT transcription.
 XX
 PS Claim 1; Page 57-58; 70pp; English.
 XX
 CC The invention provides polypeptides capable of stimulating nuclear factor
 CC (NF)-kappaB-dependent transcription or p38-dependent transcription,
 CC referred as Pellino polypeptides. The pellino polypeptides are useful for
 CC identifying modulators that alter the pellino polypeptide and pellino
 CC dominant-negative activity. They are also useful for identifying
 CC compounds that inhibit the binding activity of the polypeptides and to
 CC study cell-signal transduction. They are useful for preventing or
 CC treating infection by a pathogen such as virus, bacterial, fungi, algae
 CC or protozoa, or inhibiting apoptosis. Dominant-negative pellino
 CC polypeptides are useful for treating inflammatory conditions such as
 CC asthma, rheumatoid arthritis, inflammatory bowel disease, Crohn's
 CC disease, ulcerative colitis, atherosclerosis and Alzheimer's disease, and
 CC also for inhibiting mtogen activated protein (MAP) kinase-activated
 CC pathways. Pellino polypeptides and polynucleotides are useful to identify
 CC small molecule inhibitors of protein association or function of pellino.
 CC and other molecules involved in interleukin (IL)-1 signaling. The present
 CC sequence represents a human pellino-1 polypeptide

XX Sequence 418 AA: 83.7%; Score 1917; DB 5; Length 418;
Query Match Best Local Similarity 81.4%; Pred. No. 7,4e-188; Indels 2; Gaps 1;
Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;
QY 1 MESPQGEHCAPKPEPVKXGELVVLGYNGALPNDGRGRKSRFPALYKRPKANGVKPSTVH 60
Db 1 MESPQGEHCAPKPEPVKXGELVVLGYNGALPNDGRGRKSRFPALYKRPKANGVKPSTVH 58
QY 61 VISTPQAKAISCKGQHSISYTLSSRNOVVVEYTHDKDTDMFQVGRSTSPIDFVVTDTI 120
Db 59 IACTPQAKAISCKGQHSISYTLSSRNOVVVEYTHDKDTDMFQVGRSTSPIDFVVTDTI 118
QY 121 SGSONTDEAQITQSTISRFACRIVCORNEPYTARIAPAGDSSKNIFLEKAKAKMKNPDG 180
Db 119 PGGQNSDTSQVSTISRFACRIVCORNEPYTARIAPAGDSSKNIFLEKAKAKMKNPDG 178
QY 181 HMDGLTTNGVLVWHPRGFTESQPGVWREISVCGDVYTLRETRSAQORGLVSEETNVL 240
Db 179 QMDGLTTNGVLVWHPRGFTESQPGVWREISVCGDVYTLRETRSAQORGLVSEETNVL 238
QY 241 QDGSLLDLCGATLLMTADGLFPTTQKHEALRQETINARPCQPVGNTLAFPSINRKE 300
Db 239 QDGSLLDLCGATLLMTADGLFPTTQKHEALRQETINARPCQPVGNTLAFPSINRKE 298
QY 301 VVEKOPMAYLSCGHVGHYNNMGHRSDETEANRECEPCWCTGVGVYVPLMLGCEAGFYVDAG 360
Db 299 VVEKOPMAYLSCGHVGHYNNMGHRSDETEANRECEPCWCTGVGVYVPLMLGCEAGFYVDAG 358
QY 361 PPTHAFTPCGHVCSEKSAKYMSQIPLPHGTHAFHACPCFATQLVGEONCICKLIFQSPID 420
Db 359 PPTHAFTPCGHVCSEKSAKYMSQIPLPHGTHAFHACPCFATQLVGEONCICKLIFQSPID 418
RESULT 11
AA017518 standard; protein; 418 AA.
XX AA017518;
XX 19-JUL-2002 (first entry)
XX Human M30 variant C protein SEQ ID NO: 8.
XX Neurodegenerative disease; M30; M31; M32; M33; stroke;
XX Fragile X syndrome; Huntington's disease; Parkinson's disease;
XX Alzheimer's disease; multiple sclerosis; ovarian cancer;
XX neurodegeneration; immune disorder; autoimmune disease; allergy;
XX infection; leukaemia; inflammation; neuroprotective; cerebroprotective;
XX immunosuppressive; cytostatic; nootropic; antiparkinsonian; antiallergic;
XX virucide; antiinflammatory.
XX Homo sapiens.
XX MO200221138-A2.
XX 14-MAR-2002.
XX 07-SEP-2001; 2001WC-EP010366.
XX 07-SEP-2000; 2000US-00657479.
XX (AXAR-) AXARON BIOSCIENCE AG.
XX Schneider A, Hlemisch H, Rosner M, Klugmann M, Naim J;
XX Eisenhardt G, Kumer R, Lanahan A, Worley P, Spielvogel D, Scheek S;
XX MPI; 2002-292287/33.
XX N-PSDB; AAL46319.
XX Diagnosis of neurodegenerative disease comprises detecting level of M30-

PT family proteins.
XX Claim 2; Page 95-97; 130pp; German.
XX The present invention relates to a method of diagnosing neurodegenerative
XX diseases, comprising determining the concentration of a protein in a body
XX sample, where the protein may be M30 or a variant thereof, M31, M32 or
XX M33. The method is used to diagnose neurodegenerative diseases,
XX particularly stroke but also e.g. fragile X syndrome, Huntington's,
XX Parkinson's and Alzheimer's diseases, multiple sclerosis etc. Also
XX overexpression of M31 can be used for diagnosis of carcinoma and sarcoma,
XX especially ovarian cancer. The proteins can be used to identify specific
XX ligands, potentially useful for treating neurodegeneration, immune-system
XX disorders (e.g. autoimmune diseases, allergy, viral infection, leukaemia,
XX inflammation etc.), carcinoma and sarcoma. Inhibitors of the interaction
XX between the proteins and the protein kinase IRAK-1 can be used to treat
XX neurodegeneration. The present sequence is a protein used in the method
XX of the invention
SQ Sequence 418 AA: 83.7%; Score 1917; DB 5; Length 418;
Query Match Best Local Similarity 81.4%; Pred. No. 7,4e-188; Indels 2; Gaps 1;
Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;
QY 1 MESPQGEHCAPKPEPVKXGELVVLGYNGALPNDGRGRKSRFPALYKRPKANGVKPSTVH 60
Db 1 MESPQGEHCAPKPEPVKXGELVVLGYNGALPNDGRGRKSRFPALYKRPKANGVKPSTVH 58
QY 61 VISTPQAKAISCKGQHSISYTLSSRNOVVVEYTHDKDTDMFQVGRSTSPIDFVVTDTI 120
Db 59 IACTPQAKAISCKGQHSISYTLSSRNOVVVEYTHDKDTDMFQVGRSTSPIDFVVTDTI 118
QY 121 SGSONTDEAQITQSTISRFACRIVCORNEPYTARIAPAGDSSKNIFLEKAKAKMKNPDG 180
Db 119 PGGQNSDTSQVSTISRFACRIVCORNEPYTARIAPAGDSSKNIFLEKAKAKMKNPDG 178
QY 181 HMDGLTTNGVLVWHPRGFTESQPGVWREISVCGDVYTLRETRSAQORGLVSEETNVL 240
Db 179 QMDGLTTNGVLVWHPRGFTESQPGVWREISVCGDVYTLRETRSAQORGLVSEETNVL 238
QY 241 QDGSLLDLCGATLLMTADGLFPTTQKHEALRQETINARPCQPVGNTLAFPSINRKE 300
Db 239 QDGSLLDLCGATLLMTADGLFPTTQKHEALRQETINARPCQPVGNTLAFPSINRKE 298
QY 301 VVEKOPMAYLSCGHVGHYNNMGHRSDETEANRECEPCWCTGVGVYVPLMLGCEAGFYVDAG 360
Db 299 VVEKOPMAYLSCGHVGHYNNMGHRSDETEANRECEPCWCTGVGVYVPLMLGCEAGFYVDAG 358
QY 361 PPTHAFTPCGHVCSEKSAKYMSQIPLPHGTHAFHACPCFATQLVGEONCICKLIFQSPID 420
Db 359 PPTHAFTPCGHVCSEKSAKYMSQIPLPHGTHAFHACPCFATQLVGEONCICKLIFQSPID 418
RESULT 12
ABU00017 standard; protein; 418 AA.
XX ABU00017;
XX 17-JAN-2003 (first entry)
XX Human novel polypeptide #110.
XX Human; genetic disorder; gene mapping; medical imaging; cancer;
XX neurodegenerative disorder; lymphoid cell disorder; osteoporosis;
XX Parkinson's disease; Alzheimer's disease; bone degenerative disorder;
XX osteoarthritis; periodontal disease; liver fibrosis; viral infection;
XX fungal infection; bacterial infection; autoimmune disease; diabetes;
XX atopic dermatitis.
XX Homo sapiens.
XX OS

PN WO200274961-A1.
 XX 26-SEP-2002.
 XX 14-MAR-2002; 2002WO-US005109.
 XX 15-MAR-2001; 2001US-00810173.
 XX (HYSEQ-) HYSEQ INC.
 XX Tang YT, Zhou P, Goodrich R, Asundi V, Zhang J, Zhao QA, Ren F,
 PI Xue AJ, Yang Y, Ma Y, Yamazaki V, Chen R, Wang Z, Ghosh M,
 PI Wehrman T, Wang J, Wang D, Dmanac RT;
 XX WPI; 2003-040556/03.
 DR N-PSDB; ABX05095.
 XX
 PT New isolated polypeptides and polynucleotides, useful for preventing,
 PT treating or ameliorating medical conditions, such as cancer,
 PT neurodegenerative disorders, lymphoid cell disorders, bone degenerative
 PT disorders, and infections.
 XX
 PS Claim 9; SEQ ID NO 636; 235pp; English.
 XX
 CC The invention relates to human polynucleotides and the polypeptides they
 CC encode. The polynucleotides and polypeptides are useful in diagnostics,
 CC forensic, gene mapping, medical imaging, identification of mutations,
 CC responsible for genetic disorders or other traits, assessing biodiversity
 CC and producing many other types of data and products dependent on DNA and
 CC amino acid sequences. They are also useful for preventing, treating or
 CC ameliorating medical conditions, such as cancer, neurodegenerative
 CC disorders (e.g. Parkinson's disease, Alzheimer's disease), lymphoid cell
 CC disorders, osteoporosis, osteoarthritis, bone degenerative disorders,
 CC periodontal disease, liver fibrosis, infections (e.g. viral, fungal or
 CC bacterial) or autoimmune diseases (e.g. diabetes, atopic dermatitis).
 CC Sequences ABG9988-ABG9989 and ABU0010-ABU0043 represent human
 CC polypeptides of the invention. Note: The sequence data for this patent is
 CC not represented in the printed specification but is based on sequence
 CC information supplied by the European Patent Office
 CC
 XX
 SQ Sequence 418 AA;
 Query Match 83.7%; Score 1917; DB 6; Length 418;
 Best Local Similarity 81.4%; Pred. No. 7.4e-18;
 Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;
 QY 1 MFSPGGEHCAPNKEPVKYGELVVLGYNGALPNGDRGRKSRFPALYKRPANGVKSTVH 60
 DB 1 MFSPDQENH--PSKAPVKYGGELVVLGYNGSLPNGDRGRKSRFPALYKRPANGVKSTVH 58
 QY 61 VISTPQASKAISCKGHSISYTLSSRNOTVVEYTHDKDTDMFVGSTESPIDFVYTDI 120
 DB 59 IACTPQAKAISNKDQHSISYTLSSRAQTVEYTHDSNTDMFQIGRSTESPIDFVYTDV 118
 QY 121 SSGSNDDEAQTGSTISRFACRIVCNRPYTRIFPAGPDSKNIFLGKAAKMKPDG 180
 DB 119 PGSQNSDYSVOSTISRFACRIICERNPPTAIVAAAGPDSKNIFLGKAAKMKPSDG 178
 QY 181 HMDGLTNGVLMVMPRGFTBESQGVWRISVGVGYTLRETPSAORGVLVESENVL 240
 DB 179 QMGGLTNGVLMVMPRGFTBESQGVWRISVGVGYTLRETPSAORGVLVESENVL 238
 QY 241 QDSGLIDLGCATLLMTADGLFHTPTOKHLEALRQEIINARPOCPVGLANTLAPESINRKE 300
 DB 239 QDSGLIDLGCATLLMTADGLFHTPTOKHLEALRQEIINARPOCPVGLANTLAPESINRKE 298
 QY 301 VVEKQPMATLSCGHVGHYNNWGRSDTEANERCEPCRYGPPVPLMLGCEAGFYVDAG 360
 DB 299 VVEKQPMATLSCGHVGHYNNWGRSDTEANERCEPCRYGPPVPLMLGCEAGFYVDAG 358
 QY 361 PPTHAAPFCGVCSEKAKYWSQPLPHGTAFHFAACPPCATQLVGEONCKILFQGPID 420
 DB 359 PPTHAAPFCGVCSEKAKYWSQPLPHGTAFHFAACPPCATQLVGEONCKILFQGPID 418

RESULT 13
 ADF83103
 ID ADF83103 standard; protein; 418 AA.
 XX
 AC ADF83103;
 XX
 DT 26-FEB-2004 (first entry)
 XX
 DE Human Pellino homologue 1, overexpressed in cancer.
 XX
 KW Human; Pellino homologue 1; Pellino1; cancer; cytostatic; vaccine;
 KW gene therapy.
 XX
 OS Homo sapiens.
 XX
 PN WO2003100000-A2.
 XX
 PD 04-DEC-2003.
 XX
 XX 22-MAY-2003; 2003WO-US016049.
 XX
 XX 24-MAY-2002; 2002US-0382606P.
 XX 25-JUL-2002; 2002US-0398099P.
 XX
 XX (TULA-) TULARIK INC.
 XX
 PI Li J, Ma D, Yang J;
 XX
 DR WPI; 2004-035118/03.
 DR N-PSDB; ADF83102.
 DR GENBANK; NF_065702.
 XX
 PT Diagnosing a cancer in a mammal comprises determining RecQ5, CTXL,
 PT USP13, MCL1, or Pellino 1 gene copy number in a biological sample from a
 PT region of the mammal that is suspected to be precancerous or cancerous..
 XX
 XX
 PS Claim 179; SEQ ID NO 10; 174pp; English.
 XX
 CC The present sequence is the protein sequence of human Pellino homologue 1
 CC (Pellino1), a protein associated with the kinase domain of activated
 CC Pelle. The invention is based on the finding of the overexpression of
 CC Pellino1 and other genes (RecQ5, CTXL, USP13 and MCL1) in certain
 CC cancers, including breast cancer, colon cancer, lung cancer and ovarian
 CC cancer, and the frequent amplification of these genes in cancer cells..
 CC The genes, and their expression products, can be used diagnostically or
 CC as targets for cancer therapy. They can also be used to identify and
 CC design compounds useful in the diagnosis, prevention and therapy of
 CC tumours and cancers, in vaccine development, and in methods for
 CC determining the efficacy of a treatment regime. A claimed method for
 CC inhibiting cancer or precancerous growth, especially in ovarian or lung
 CC tissue, uses an inhibitor that interacts with Pellino1 DNA or RNA. The
 CC inhibitor is a small interfering RNA (siRNA), microRNA (miRNA), an
 CC antisense RNA, and antisense DNA, a decoy molecule, a decoy DNA, a
 CC ribozyme or small molecule.
 CC
 XX
 SQ Sequence 418 AA;
 Query Match 83.7%; Score 1917; DB 8; Length 418;
 Best Local Similarity 81.4%; Pred. No. 7.4e-18;
 Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;
 QY 1 MFSPGGEHCAPNKEPVKYGELVVLGYNGALPNGDRGRKSRFPALYKRPANGVKSTVH 60
 DB 1 MFSPDQENH--PSKAPVKYGGELVVLGYNGSLPNGDRGRKSRFPALYKRPANGVKSTVH 58
 QY 61 VISTPQASKAISCKGHSISYTLSSRNOTVVEYTHDKDTDMFVGSTESPIDFVYTDI 120
 DB 59 IACTPQAKAISNKDQHSISYTLSSRAQTVEYTHDSNTDMFQIGRSTESPIDFVYTDV 118
 QY 121 SSGSNDDEAQTGSTISRFACRIVCNRPYTRIFPAGPDSKNIFLGKAAKMKPDG 180

Db 119 PGSQNSDTSOSTISRFACRIICERNPFTARIYAAGDSSKNIFLGEKAKMKTSDG 178
 Qy 181 HMDGLTTNGVLYVNHPRGFTBESOPGVWREISVCGDYTLRETRSAOQRKLVESNTNV 240
 Db 179 QMDGLTTNGVLYVNHPRGFTBESOPGVWREISVCGDYTLRETRSAOQRKLVESNTNV 238
 Qy 241 QDGSLLDLCGATLLMTADGLFHTPTQKHEALROEINARPOCPVGLNTLAFPSINRKE 300
 Db 239 QDGSLLDLCGATLLMTADGLFHTPTQKHEALROEINARPOCPVGLNTLAFPSINRKE 298
 Qy 301 VVEKOPWATLSCGHVGHYNNMGRSDTEANERECPCRTVGPVPLMTGCEAGFYVDAG 360
 Db 299 VDEKOPWVYLVNCGHVGYNMKNKEDRCKRCEPCRSVGPVPLMTGCEAGFYVDAG 358
 Qy 361 PPTHAFPCGHCVCSEKSAKTYWSQIPLPHGTHAFHAACPCATQVGEONCICKLIFQSPID 420
 Db 359 PPTHAFSPCGHVCSEKTYATWSQIPLPHGTHAFHAACPCATQVGEONCICKLIFQSPID 418
 RESULT 14
 ID ADO44005 standard; protein; 418 AA.
 AC ADO44005;
 DT 15-JUL-2004 (first entry)
 XX
 DE Amino acid sequence of human Pellino 1.
 XX
 protein complex; neurological disease; stroke; neurodegeneration;
 KM Wallerian degeneration; Alzheimer's disease; neurological disorder;
 KM epilepsy; inflammatory condition; ulcerative colitis; Crohn's disease;
 KM atherosclerosis; ID-MYO-inositol triphosphate 3 kinase A; ASK1; ASK2;
 KM ASK3; Cdk37; CHK2; CTCL tumour antigen SE20-4; EF-1 alpha 1; EMAP;
 KM Cdc37; CHK2; CTCL tumour antigen SE20-4; EF-1 alpha 1; EMAP;
 KM FLJ14653 NTRP2002252; FLJ30839 FEBRA2002429; HERC2;
 KM inositol polyphosphate-5-phosphatase; inositol-1; 4;
 KM 5-triphosphate 5-phosphatase type 1; IRAK1; IRAK4; KIAA1441; MSTP030;
 KM Nek9; PA3; Pellino 1; Pellino 3; podocalyxin-like protein 1 precursor;
 KM Pushover; S-adenosylhomocysteine;
 KM secretory carrier-associated membrane protein 2; surfactant protein 2;
 KM ubiquitin carboxyl terminal hydrolase 11;
 KM upstream regulatory element binding protein 1; Vartul;
 KM Werner's syndrome helicase interacting protein; WHIP;
 KM X-ray repair cross complementing protein 4.
 XX
 OS Homo sapiens.
 PN MO2004031242-A2.
 PD 15-APR-2004.
 XX
 PF 11-SEP-2003; 2003MO-EP010110.
 XX
 PR 12-SEP-2002; 2002BP-00020495.
 PR 12-SEP-2002; 2002BP-00020496.
 PR 12-SEP-2002; 2002EP-00020497.
 XX
 PA (CELL-) CELLZONE AG.
 PI Boummeester T, Drewes G, Jackson D, Helftenbein G, Schirle M;
 PI Kuester B, Hopf C;
 XX
 DR WPI; 2004-316467/29.
 XX
 PT New complex comprising at least one first protein, and at least one
 PT second protein, useful for treating stroke, Alzheimer's disease,
 PT neurological disorders such as epilepsy, and inflammatory conditions such
 PT as ulcerative colitis.
 XX
 PS Example; Page 224-226; 287pp; English.
 XX
 CC The specification describes protein complexes involved in cellular

CC processes which have been shown to be critical for the development of
 CC various forms of neurological diseases. Three protein complexes were
 CC identified: ASK2 protein complex, Pellino-1 protein complex and Pellino-3
 CC protein complex. The protein complex are useful for treating diseases and
 CC disorders, e.g. stroke, neurodegeneration such as Wallerian degeneration,
 CC Alzheimer's disease, neurological disorders such as epilepsy, and
 CC inflammatory conditions such as ulcerative colitis, Crohn's disease or
 CC atherosclerosis. Proteins identified as being part of the protein
 CC complexes of the invention are ID-MYO-inositol triphosphate 3 kinase A,
 CC ASK1, ASK2, ASK3, Cdk37, CHK2, CTCL tumour antigen SE20-4, EF-1 alpha 1, EMAP,
 CC FLJ14653 NTRP2002252, FLJ30839 FEBRA2002429, HERC2, two hypothetical
 CC proteins of 35.5 kDa and 49.3 kDa, inositol polyphosphate-5-phosphatase,
 CC inositol-1,4,5-triphosphate 5-phosphatase type 1, IRAK1, IRAK4, KIAA1441,
 CC MSTP030, Nek9, PA3, Pellino 1, Pellino 3, podocalyxin-like protein 1
 CC precursor, Pushover, a putative S-adenosylhomocysteine, secretory
 CC carrier-associated membrane protein 2, surfactant protein 2, ubiquitin
 CC carboxyl terminal hydrolase 11, upstream regulatory element binding
 CC protein 1, Vartul, Werner's syndrome helicase interacting protein (WHIP),
 CC X-ray repair cross complementing protein 4 (isoform 1). The present
 CC sequence represents Pellino 1.
 CC
 SQ Sequence 418 AA;
 Query Match 83.7%; Score 1917; DB 8; Length 418;
 Best Local Similarity 81.4%; Pred. No. 7, 4e-188;
 Matches 342; Conservative 38; Mismatches 38; Indels 2; Gaps 1;
 Qy 1 MSPGQEBHCAPKKEVYKIGELVYLGNALPNGBDGRKRSRALKPRKANGKSTYH 60
 Db 1 MSPDDEBNH--PEKAVKVGELVLVLYGNSLPNGDGRKRSRALKPRKANGKSTYH 58
 Qy 61 VISTPQASAKISCKGSHSISTYLSRNQTVVEETHDKDDMPQVGRSTSPIDFVYTDIT 120
 Db 59 IACTPQAKAISCKKHDSISTYLSRNQTVVEETHDKDDMPQVGRSTSPIDFVYTDIT 118
 Qy 121 SSGSNTDEAQTOSTISRFACRIICERNPFTARIYAAGDSSKNIFLGEKAKMKTSDG 180
 Db 119 PGSQNSDTSOSTISRFACRIICERNPFTARIYAAGDSSKNIFLGEKAKMKTSDG 178
 Qy 181 HMDGLTTNGVLYVNHPRGFTBESOPGVWREISVCGDYTLRETRSAOQRKLVESNTNV 240
 Db 179 QMDGLTTNGVLYVNHPRGFTBESOPGVWREISVCGDYTLRETRSAOQRKLVESNTNV 238
 Qy 241 QDGSLLDLCGATLLMTADGLFHTPTQKHEALROEINARPOCPVGLNTLAFPSINRKE 300
 Db 239 QDGSLLDLCGATLLMTADGLFHTPTQKHEALROEINARPOCPVGLNTLAFPSINRKE 298
 Qy 301 VVEKOPWATLSCGHVGHYNNMGRSDTEANERECPCRTVGPVPLMTGCEAGFYVDAG 360
 Db 299 VDEKOPWVYLVNCGHVGYNMKNKEDRCKRCEPCRSVGPVPLMTGCEAGFYVDAG 358
 Qy 361 PPTHAFPCGHCVCSEKSAKTYWSQIPLPHGTHAFHAACPCATQVGEONCICKLIFQSPID 420
 Db 359 PPTHAFSPCGHVCSEKTYATWSQIPLPHGTHAFHAACPCATQVGEONCICKLIFQSPID 418
 RESULT 15
 ID ADP48667 standard; protein; 418 AA.
 AC ADP48667;
 DT 09-SEP-2004 (first entry)
 XX
 DE Human Pellino-1 protein SEQ ID NO:4.
 XX
 KM nuclear factor kappa B dependent transcription inhibitor;
 KM NF-kB-dependent transcription inhibitor; Pellino-1;
 KM p38-dependent transcription inhibitor; Pellino-1;
 KM interleukin 1 receptor-associated kinase 4;
 KM IL-1 receptor-associated kinase 4; IRAK-4; antiarteriosclerotic;
 KM antiarthritic; antiasthmatic; antiinflammatory; antineumatic; antiulcer;

This Page Blank (uspto)